DISRUPTING THE QUATERNARY STRUCTURE OF DHDPS AS A NEW APPROACH TO ANTIBIOTIC DESIGN

A thesis submitted in partial fulfilment of the requirements
for the degree of

Doctor of Philosophy in Biochemistry

at the

University of Canterbury

by

Genevieve L. Evans

2009

To Jesus who anchored me

Among scientist are collectors, classifiers, and compulsive tidy uppers; many are detectives by temperament and many are explorers; some are artists and others artisans. There are poet scientist and philosopher scientist and even a few mystics.

-- Peter Medawar, from *The strange case of the spotted mice*

Acknowledgements

One page is not enough to express the many thanks I owe to all the people who have encouraged and supported me, both within and outside the scientific field over the course of this PhD. Special thanks go to my dear friends, Audrey and Jennifer, who loved me when I was the craziest, most stressed-out version of myself. “I thank my God upon every remembrance of you” - Philippians 1:3

This PhD would not have been possible without my supervisor, Professor Juliet Gerrard, who has been a constant spring of encouragement and wisdom. Thank you for your patience. It has been a privilege to have a supervisor who wants to grow scientists as well as do great science.

Thank you to all the past and present members of the purple bubble and team DHDPS who have provided help and input over the years. It has been an adventure that would have been much harder without your kindness. I am especially grateful for the encouragement and comradeship from those who battled beside me in the trenches of lab work towards achieving a Postgrad degree: Megan, Pam, Shiva, Sophie, Tatiana and Voula-“watch-out-for-those-negatrons”-Mitsakos. Thanks also to Jackie Healey for being a constant and steadying influence in the lab, and putting up with lots of notes scribbled on sturdy lab-quality paper towel.

Many thanks to the four musketeers, Ren, Sean, Andy and Grant, such gifted scientists, who were generous enough to answer my unending barrage of questions, when I finally found the courage to ask them. Special thanks go to Sean for proof-reading this thesis, attempting to make it more concise and to sort out my “affects” from “effects”.

Big thanks go to Dr Matt Perugini and Dr Mike Griffin at Bio21 (Melbourne), for making time to share enthusiasm and practical know-how with regards to the exciting technique of analytical ultracentrifugation. Thanks to Mike, especially, for enabling the collection of some quality data and providing superior cross-Tasman support when it came to data processing.

Big thanks also go to Professor Geoff Jameson at Massey University, who generously gave time and guidance in structural refinement, thesis drafts, and more, including the most

effective definitions of “affect and “effect” that has forever affected (and hopefully improve) my usage of the words.

Much appreciation is extended to Dr Manfred Weiss, at the EMBL Hamburg Outstation, and Dr Georgia Kefala, who provided the plasmid which made this work possible, and Ren for forging this fruitful collaboration. Much thanks also go to Manfred’s PhD student, Linda Schultz for her friendly emails and for crystallized my mutant enzyme, one of the crowning achievements in this PhD.

Thanks to my Dad for always encouraging me to seek understanding rather than memorization, because understanding is more important. Thanks to my Mom for imparting her sense of beauty and curiosity about the way things work in nature. Thanks to my siblings Rosalie and Benjamin for the parts of this journey you’ve shared with me.

Finally and most importantly, I thank God, for making a world as beautiful and intricately complicated as to contain enzymes such as DHDPS, and giving me the curiosity to want to know how they work.

Table of Contents

APPENDICES

Abbreviations

A

Å	Angstrom
A / ALA	alanine (amino acid)
A	adenine (nucleotide)
Amp	ampicillin
Amp^r	ampicillin resistance
(*S*)-ASA	(*S*)-aspartic β-semialdehyde
AU	absorbance units
AUC	analytical ultracentrifugation

B

β-ME	β-mercaptoethanol
BCG	bacillus Calmette-Guerin vaccine, derived from *Mycobacterium bovis* by Calmette and Guérin
Bicine	2-(Bis(2-hydroxyethyl)amino)acetic acid
BN-PAGE	blue native polyacrylamide gel electrophoresis
bp	base pair
BSA	bovine serum albumin (protein)

C

χ^2	chi-squared (statistical measure of global fit)
ºC	degrees Celsius
C / CYS	cysteine (amino acid)
C	cytosine (nucleotide)
Cα	α-carbon atom
Cam	chloramphenicol
Cam^r	chloramphenicol resistance
CD	circular dichroism
C. glutamicum	*Corynebacterium glutamicum*
cP	centipoise
[*c(s)*]	continuous sedimentation coefficient distribution
[*c(M)*]	continuous molar mass distribution
cm	centimetre

CuZn-SOD	copper-zinc superoxide (protein)
CV	column volumes

D

D / ASP	aspartate (amino acid)
Da	Dalton
DAP	*meso*-diaminopimelate (amino acid)
*dap*A	gene encoding dihydrodipicolinate synthase
*dap*B	gene encoding dihydrodipicolinate reductase
DAH7PS	3-deoxy-D-arabino-heptulosonate-7-phosphate synthase (protein)
DHDP	dihydrodipicolinate
DHDPS	dihydrodipicolinate synthase (protein)
DHDPR	dihydrodipicolinate reductase (protein)
d-H_2O	distilled water
DMF	dimethylformamide
DNA	deoxyribonucleic acid
DNAse	deoxyribonuclease, degrades DNA (protein)
dNTP	deoxyribonucleotide; *i.e.* single unit of DNA
Dpn I	restriction enzyme; an endonuclease specific for methylated & hemimethylated DNA (protein)
DSF	differential scanning fluorescence
DTT	dithiothreitol

E

ε	molar extinction coefficient
E / GLU	glutamate (amino acid)
e_0	enzyme molarity (MW expressed in units of mg.μmol), used to convert V to k_{cat}
EB (buffer)	Tris.HCl buffer for storing DNA
E. coli	*Escherichia coli*
EDTA	ethylene diamine tetraacetic acid
*end*A	gene encoding endonuclease I
ESI	electrospray ionization
EtOH	ethanol

F

f	frictional coefficient
F / PHE	phenylalanine (amino acid)
F	farad; SI unit of capacitance
F	geometric factor
f/f_0	frictional ratio
f_m	fraction of monomer

G

g	gram
G / GLY	glycine (amino acid)
G	guanine (nucleotide)
groEL	gene encoding GroEL chaperonin
groES	gene encoding GroES chaperonin

H

η	solvent viscosity
H / HIS	histidine (amino acid)
HCl	hydrochloric acid
HEPES	*N*-2-hydroxyethylpiperazine-*N*′-2-ethane sulfonic acid
His_6-tag	peptide purification tag containing 6 histidine residues
HIV	human immunideficiency virus
HPLC	high-pressure liquid-chromatography
hrs	hours
HTPA	(4*S*)-4-hydroxy-2,3,4,5-tetrahydro-(2*S*)-dipicolinic acid

I

I / ISO	isoleucine (amino acid)
IC_{50}	inhibitor concentration resulting in 50% inhibition
IPTG	isopropyl-1-thio-β-D-galactoside
IS	ionic strength

J

JPD	JavaProtein Dossier

K

K / LYS	lysine (amino acid)
Kan	kanamycin
Kan^r	kanamycin resistance
kcal	kilocalorie
k_{cat}	catalytic turnover-number
k_{cat} / K_M	catalytic efficiency
kDa	kiloDalton
$K_D^{2\rightarrow1}$	dimerization dissociation constant
KDGA	2-keto-3-deoxygluconate aldolase
K_i	dissociation constant that defines the strength of inhibition
K_M	Michaelis constant
kV	kilovolt

L

L / LEU	leucine (amino acid)
L	litre
lac	operon encoding β-galactosidase
LB	Luria-Bertani medium
LCMS	liquid-chromatography mass spectrometry
LDH	lactate dehydrogenase (protein)
LDS	lithium dodecyl sulfate
*lys*A	gene encoding DAP decarboxylase

M

µL	microlitre
µF	microfarad
µg	microgram
µL	microlitre
µm	micrometre
µmol	micromole
µM	micromolar
M / MET	methionine (amino acid)
M	moles per litre, *i.e.* molar

M	molecular weight
M. bovis	*Mycobacterium bovis*
MDH	malate dehydrogenase (protein)
MDR-TB	multi-drug resistant strain of *M. tuberculosis*
MeOH	methanol
MES	2-(*N*-morpholino)ethanesulfonic acid
mg	milligram
min	minute
M. jannaschii	*Methanocaldococcus jannaschii*
mL	millilitre
M. leprae	*Mycobacterium leprae*, causative agent of leprosy
mM	millimolar
MOPS	3-morpholinopropane-1-sulfonic acid
MRSA	methicillin-resistant *Staphylococcus aureus*
M. smegmatis	*Mycobacterium smegmatis*
M. tuberculosis	*Mycobacterium tuberculosis*, causative agent of tuberculosis
MW	molecular weight
mV	millivolt

N

N / ASN	asparagine (amino acid)
N	Avogadro's number
NaCl	sodium chloride (salt)
NADP+	nicotamide adenine dinucleotide phosphate, oxidized form
NADPH	nicotamide adenine dinucleotide phosphate, reduced form
NAL	N-acetylneuraminate lysase (protein)
NaOH	sodium hydroxide (base)
NCS	non-crystallographic similarity
N. sylvestris	*Nicotiana sylvestris* (tobacco)
ng	nanogram
nL	nanolitre
nm	nanometre
nM	nanomolar
NMR	nuclear magnetic resonance

NOS	nitric oxide synthase (protein)

O

OD_{600}	optical density *i.e.* absorbance at 600 nm

P

π-(interactions)	interactions between the π-orbitals of aromatic residues
P / PRO	proline (amino acid)
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDB	Protein Data Bank
pers. comm.	personal communication
pH	$-\log_{10} [H^+]$
PISA	Protein Interfaces, Surfaces, and Assemblies database
P_T	total protein concentration

Q

Q / GLN	glutamine (amino acid)

R

ρ	solvent density
R / ARG	arginine (amino acid)
r	radial distance from the centre of rotation
R-(factor)	refinement value
R^2	coefficient of determination (statistical measure of fit)
*rec*A	recombination deficient
r.m.s.d.	root-mean-square difference
RNA	ribonucleic acid
RNAse	ribonuclease, degrades RNA (protein)
rpm	revolutions per min

S

s	second
s	sedimentation coefficient
s^*	observed sedimentation coefficient
$s°_{20,w}$	standardized sedimentation coefficient
S / SER	serine (amino acid)

S	Svedberg (unit)
[*S*]	substrate concentration
SCS	succinyl-CoA synthetase
SDS	sodium dodecyl sulfate
SDS-PAGE	denaturing polyacrylamide gel electrophoresis
SE	sedimentation equilibrium
SOB	super optimal broth
SOC	super optimal broth, with catabolite repression
SV	sedimentation velocity

T

T / THR	threonine (amino acid)
T	thymine (nucleotide)
TAE	Tris.acetyl EDTA electrophoresis buffer
TCA	trichloroacetic acid
TE (buffer)	Tris.HCl EDTA buffer for storing DNA
Tet	tetracycline
Tet^r	tetracycline resistance
TEV	tobacco etch virus
TI (noise)	radial-dependent, time invariant error due to optical imperfections in cell used in SV experiments
TIM	triosephosphate isomerase (protein)
TLS	translation, rotation, screw-rotation
T_m	melting temperature
T_m^{app}	apparent melting temperature, *i.e.* mid-point temperature between folded and unfolded states
T. maritima	*Thermotoga maritima*
TNF	tumour-necrosis factor (protein)
tRNA	transfer RNA
Tris.acetate	tris(hydroxymethyl) aminomethane, adjust to pH with glacial acetic acid
Tris.base	tris(hydroxymethyl) aminomethane
Tris.HCl	tris(hydroxymethyl)aminomethane hydrochloride, adjust to pH with HCl

U

u	rate of sedimentation

UV	ultraviolet

V

ν	initial-rate of enzyme activity, measured by the consumption of NADPH
$\bar{v}$	partial specific volume
V / VAL	valine (amino acid)
V	maximum velocity or rate of enzyme activity
V	volt; SI unit of electric potential
V_e	elution volume
v/v	describing in terms of volume per volume, *e.g.* litre per litre

W

ω	angular velocity (radians/seconds)
Ω	ohm, unit of electrical resistance
W / TRY	tryptophan (amino acid)
w/v	describing in terms of mass per volume, *e.g.* gram per millilitre
WHO	World Health Organization

X

XDR-TB	extensively drug resistant strain of *M. tuberculosis*
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Y

Y / TYR	tyrosine (amino acid)

Z

Z-test	statistical test of fit

Abstract

This thesis examined the enzyme dihydrodipicolinate synthase (DHDPS, E.C. 4.2.1.52) from the pathogen *Mycobacterium tuberculosis*. DHDPS is a validated antibiotic target for which no potent inhibitor based on substrates, intermediates or product has been found. The importance of the homotetrameric quaternary structure in *E. coli* DHDPS has been demonstrated by the 100-fold decrease in activity observed in a dimeric variant, DHDPS-L197Y, created by site-directed mutagenesis. This suggested a new approach for inhibitor design: targeting the dimer-dimer interface and disrupting tetramer formation.

DHDPS catalyzes the first committed step in the biosynthetic pathway of *meso*-diaminopimelic acid, a critical component of the mycobacterial cell wall. In this study, wild-type *M. tuberculosis* DHDPS was thoroughly characterized and compared with the *E. coli* enzyme. A coupled assay was used to obtain the kinetic parameters for *M. tuberculosis* DHDPS: $K_M^{(S)\text{-}ASA}$ = 0.43 (±0.02) mM, $K_M^{pyruvate}$ = 0.17 (±0.01) mM, and k_{cat} = 138 (±2) s^{-1}. Biophysical techniques showed *M. tuberculosis* DHDPS to exist as a tetramer in solution. This is consistent with the crystal structure deposited as PDB entry 1XXX. The crystal structure of *M. tuberculosis* DHDPS showed active-site architecture analogous to *E. coli* DHDPS and a dimeric variant of *M. tuberculosis* DHDPS was predicted to have reduced enzyme activity.

A dimeric variant of *M. tuberculosis* DHDPS was engineered through a rationally designed mutation to analyze the effect of disrupting quaternary structure on enzyme function. A single point mutation resulted in a variant, DHDPS-A204R, with disrupted quaternary structure, as determined by analytical ultracentrifugation and gel-filtration chromatography. DHDPS-A204R was found to exist in a concentration-dependent monomer-dimer equilibrium, shifted towards dimer by the presence of pyruvate, the first substrate that binds to the enzyme. The secondary and tertiary structure of DHDPS-A204R was analogous to wild-type *M. tuberculosis* DHDPS as judged by circular dichroism spectroscopy and X-ray crystallography, respectively. Surprisingly, this disrupted interface mutant had similar activity to the wild-type enzyme, with a k_{cat} of 119 (±6) s^{-1}; although, the affinity for its substrates were decreased: $K_M^{(S)\text{-}ASA}$ = 1.1 (±0.1) mM, $K_M^{pyruvate}$ = 0.33 (±0.03) mM. These results indicated that disruption of tetramer formation does not provide an alternative direction for drug design for DHDPS from *M. tuberculosis*.

Comparison with the recently discovered dimeric DHDPS from *Staphylococcus aureus* shed further light on the role of quaternary structure in DHDPS. In *M. tuberculosis* DHDPS-A204R and the naturally dimeric enzyme, the association of monomers into the dimer involves a greater buried surface area and number of residues than found in *E. coli* DHDPS-L197Y. This provides a framework to discriminate which DHDPS enzymes are likely to be inactive as dimers and will direct future work targeting the dimer-dimer interface of DHDPS as an approach for drug design.

Chapter One

Introduction

The subject of this thesis is the tetrameric enzyme dihydrodipicolinate synthase (DHDPS) from *Mycobacterium tuberculosis.* This enzyme catalyzes the first committed step of the (*S*)-lysine biosynthetic pathway, which also produces *meso*-diaminopimelic acid (DAP). Both (*S*)-lysine and DAP are essential components of the bacterial cell wall, and the biosynthetic pathway is absent in humans, which makes DHDPS an attractive target for drug design. Several genomic studies have classed DHDPS as the product of an essential gene, and thus validated it as a target for antibiotic design.[1]

This work explores the connection between quaternary structure and enzyme function. The importance of the tetrameric structure of *Escherichia coli* DHDPS has previously been demonstrated: a dimeric *E. coli* DHDPS, created by mutating one of the residues involved in inter-subunit interactions, showed significantly less activity than the tetrameric wild-type enzyme.[2] If this proves to be a general observation for DHDPS enzymes from pathogenic micro-organisms, it could open a new avenue of antibiotic design, as molecules could be developed to break up the tetrameric structure and deactivate DHDPS. Therefore, the aim of this thesis is to examine the quaternary structure of DHDPS from *M. tuberculosis* as a putative drug target.

1.1 Tuberculosis

There is a desperate need for new drugs to treat *M. tuberculosis* infection - it has been 50 years since a new anti-tuberculosis drug has been approved for treatment.[3] *M. tuberculosis* causes more death in the world today than any other bacterial agent,[4] and the death toll is set to rise as drug-resistant strains emerge.[5] In 2005, 1.6 million people died from tuberculosis and it is a leading killer among HIV-infected people with weakened immune systems.[6]

Within a year of the discovery of the first antibiotic against *M. tuberculosis*, in 1944, there were already resistant strains emerging.[7] This pattern has continued and consequently current treatment involves a multi-drug approach. The symptoms of tuberculosis tend to disappear

within the first month of treatment, in spite of the infection still being present, so patients often fail to complete an adequate course of antibiotics, which increases the likelihood of drug resistance arising.[4] The adaptability of *M. tuberculosis* is alarming and, recently, the World Health Organization (WHO) has defined a new kind of drug resistant strain that is not only resistant to almost all of the commonly used, so-called "first-line", drugs but also several of the more costly second-line drugs, thus almost eliminating all methods of treatment.[6] In a case study in South Africa, 52 patients were identified to be infected with such a strain, and all but one died as a result.[8] It is believed that this extensively drug resistant strain, XDR-TB, has developed from MDR-TB, a multi-drug resistant strain which shows resistance to several first-line drugs.[3] The WHO survey of 76 countries, from 1999-2002, showed 0 to 14 % (median 1 %) of new tuberculosis infections were caused by MDR-TB.[9]

The first vaccine against *M. tuberculosis*, BCG (bacillus Calmette-Guerin), derived from *M. bovis* by Calmette and Guérin, has been distributed since the 1920s and remains the only available vaccine for humans.[10] Unfortunately, there is significant regional variation in the protection provided by the BCG vaccine against pulmonary tuberculosis, the most common form of tuberculosis which occurs in the lungs, with reported variation in efficacy between 0 and 80 %.[11] However, the BCG vaccine is effective against the childhood manifestations of *M. tuberculosis*, miliary or meningeal tuberculosis, occurring in the blood and brain, respectively, and its close relative, *M. leprae*, the causative agent in leprosy, and therefore the vaccine continues to be used extensively.[10] Consequently, rather than looking for a replacement vaccine, much of the current research is exploring the development of a "booster" vaccine, from adjuvants containing proteins, or proteins fused to DNA.[12]

The complete sequencing of the *M. tuberculosis* strain H37Rv genome is probably the most important advance in the mycobacterial field since the identification of the bacillus,[13] by R. Koch a century ago.[4] More than 3900 genes thought to encode proteins were identified and at least 52 % of these have been assigned function.[14,15] Additionally, several other strains of *M. tuberculosis* have been sequenced and annotated (H37Ra,[16] CDC1551,[17] and F11[18]). No variation in the amino-acid sequence for DHDPS has been observed, suggesting that the enzyme is not linked to the variation in virulence and drug susceptibility in these strains. *M. bovis*, the most closely related species to *M. tuberculosis,* also has an identical sequence for DHDPS but other relatives, such as *M. leprae*, show some differences (87 % identity determined using CLUSTALW).[19]

The conservation of a gene encoding DHDPS, especially in *M. leprae*, described as the minimal mycobacterial gene set,[20] provides evidence for the essential nature of the enzyme DHDPS. Computer simulations suggest 35 % of the *M. tuberculosis* genome is essential;[21] however, *in vitro* experiments have only found 614 genes necessary for optimal growth, including the gene coding for DHDPS.[22] Most of the enzymes of the (*S*)-lysine biosynthetic pathway have been shown to be necessary or essential for growth under normal conditions in various bacterial genome projects.[1]

1.2 The metabolic role of DHDPS

The enzyme DHDPS catalyzes the condensation of (*S*)-aspartate semialdehyde ((*S*)-ASA) and pyruvate in the first committed step of the (*S*)-lysine/DAP biosynthetic pathway.[23] The substrate (*S*)-ASA is derived from (*S*)-aspartic acid, and can alternatively be reduced by homoserine dehydrogenase to give (*S*)-homoserine, a precursor to (*S*)-methionine, (*S*)-threonine and (*S*)-isoleucine **(Figure 1.1)**.[24] These four amino acids, together with (*S*)-lysine and DAP, comprise the aspartic acid family of amino acids, and homoserine dehydrogenase, together with DHDPS, have key regulator roles as branch point enzymes.[25]

1.2.1 The role of DHDPS in the DAP pathway

There are four variants of the DAP/(*S*)-lysine biosynthetic pathway found in bacteria **(Figure 1.1)** which use slightly different sets of enzymes.[26-28] The acetylase pathway appears to be limited to certain *Bacillus* species,[26] with the notable exception of *Bacillus sphaericus* which utilizes the dehydrogenase pathway instead.[29] Mycobacteria, like *E. coli* and other Gram-negative bacteria, use the succinylase pathway.[30] The aminotransferase pathway has only recently been discovered in the bacterial genus *Chlamydia*, which lack a detectable peptidoglycan cell wall, and there is growing evidence of its use in plants.[27] Comparative genomic analysis, coupled with experimental verification, have determined further lineages of eubacteria, as well as two archaeal groups, which utilize the aminotransferase pathway.[31] Several of these bacteria are suggested to use two of the four pathways; that is, the dehydrogenase pathway in addition to aminotransferase pathway,[31] in a similar manner to the Gram-positive bacteria, *Corynebacterium glutamicum*, which uses both the succinylase and dehydrogenase pathway.[32] Consequently, any antibacterial agent used against these bacteria would have to inhibit an enzyme common to both pathways, such as DHDPS.

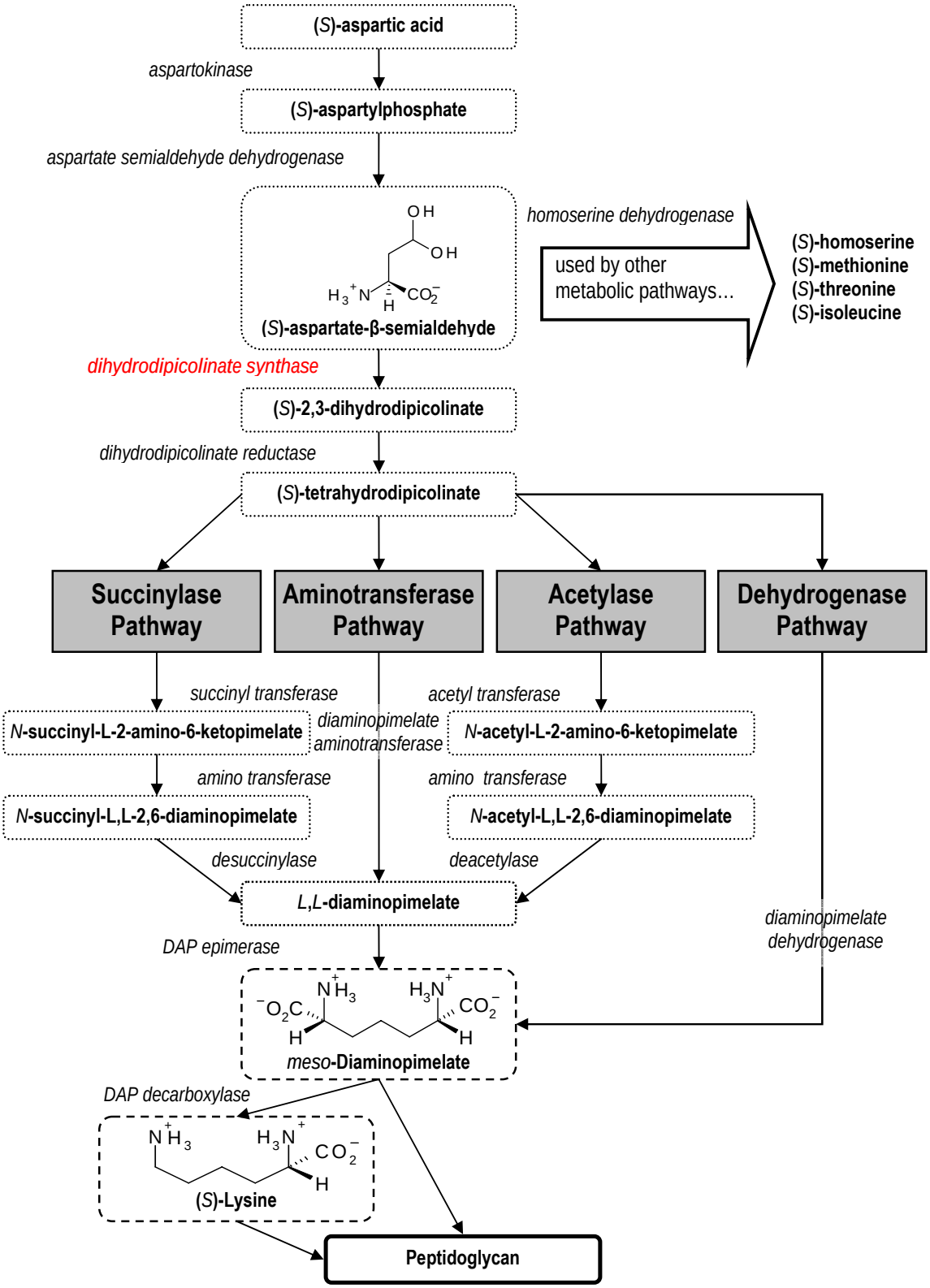


Figure 1.1: A schematic diagram of the four known biosynthetic pathways utilized by bacteria to produce DAP and (*S*)-lysine, which are used as components in the peptidoglycan cell wall.[26-28]

All these pathways share initial and terminal steps, so the enzymes catalyzing these steps, which include DHDPS, dihydrodipicolinate reductase (DHDPR), and DAP decarboxylase, should be conserved in all bacterial species.[27,31] DHDPS being crucial in all four pathways, therefore, provides an opportunity for developing broad spectrum drugs.[33]

1.2.2 DAP/(*S*)-lysine biosynthesis as a target for antibacterial agents

First discovered by Work in 1950 in *Corynebacterium diphtheriae*,[34] and subsequently in *M. tuberculosis*,[35,36] DAP is the precursor to an essential amino acid, (*S*)-lysine, and also has a role unique to bacteria.[37,38] In almost all bacteria, either DAP or (*S*)-lysine are part of the short peptides that cross-link the long polymers composing the cell wall.[39] They play a crucial role linking the tetrapeptide branches **(Figure 1.2)** that give the peptidoglycan cell wall the strength to withstand high intracellular osmotic pressure, and consequently in bacteria the absence of (*S*)-lysine or DAP should result in cell lysis and death.[39]

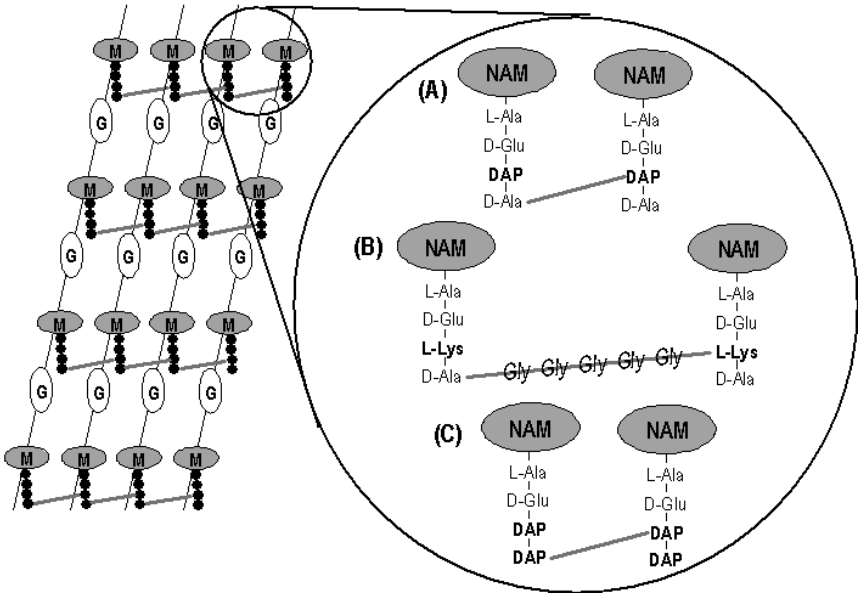


Figure 1.2: **A schematic diagram of peptidoglycan, which is composed primarily of long alternating sequences of *N*-acetylglucosamine (G) and *N*-acetylmuramic acid (M/NAM) interlinked by peptide chains. The peptide cross-link sequences vary; (A) is mostly found in Gram-negative bacteria, (B) in Gram-positive bacteria,[26] and (C) is unusual and found in mycobacteria.[40] The tetrapeptide branches are bonded directly to each other in Gram-negative bacteria and mycobaceria, whereas an additional sequence of five glycines (Gly) is required to make the link in Gram-positive bacteria.**

Many existing antibiotics inhibit other key steps in the assembly of the bacterial cell wall and the biosynthesis of its components has long been accepted as a target for antibiotic design.[41] An inhibitor which inactivated one of the enzymes in the DAP/(*S*)-lysine biosynthetic pathway and prevented the synthesis of these crucial amino acids would be a very effective antibiotic.[1,26,38,39,41,42] These antimicrobial molecules would constitute a new class, as no naturally occurring antibiotics have been found that act on the DAP biosynthetic pathway.[41]

Mycobacteria have cell walls with an unusually high DAP content because they contain an unusual tetrapeptide linkage, where the second alanine (D-Ala) in the peptide sequence has been replaced with DAP **(Figure 1.2C)**.[40] The resulting DAP-DAP link within the tetrapeptide constitutes one third of the cell wall interpeptide linkages in *M. smegmatis,*[40] and half of those in *M. bovis,*[43] and elevates the importance of DAP in the mycobacterium cell wall.[40] This might mean that bacteria from this classification are more sensitive to compounds that inhibit the DAP/(*S*)-lysine biosynthetic pathway. DAP auxotrophs have been created from several bacterial species, including *E. coli* and *M. smegmatis*, which in the absence of DAP undergo cell lysis and death.[25] These gene-knockout experiments demonstrate the essential nature of DAP and validate the DAP biosynthetic pathway as a target for antibacterial drug design.

Unfortunately, it was also found that *M. smegmatis* DAP auxotrophs spontaneously mutated to suppress their requirement for DAP,[44] using the bacterial metabolite lanthionine as an alternative cell wall component.[45] This provides a possible resistance mechanism against antibiotics that inhibit the DAP biosynthetic pathway. Interestingly, these mutants were hypersensitive to β-lactam antibiotics,[44] reflecting the altered cell wall structure.[45] While the exact role of the DAP-DAP linkage is unknown, it has been linked to mycobacterial resistance to β-lactam antibiotics.[46] Developing synergistic combinations of antibiotics has been suggested as a means for overcoming β-lactam antibiotic resistance,[47] and like β-lactamase inhibitors,[48] a drug that inhibits DAP synthesis, could be used in combination with β-lactams to counteract antibiotic resistance.

Disrupting (*S*)-lysine production is another rationale for targeting the enzymes of the DAP biosynthetic pathway. Unlike DAP, (*S*)-lysine is an essential amino acid for protein synthesis. A (*S*)-lysine biosynthetic pathway, involving DAP as an intermediate, is also found in plants.[41] An unrelated (*S*)-lysine pathway, named the α-aminoadipate pathway, is found in

higher fungi and euglenoids.[49] Mammals also require (*S*)-lysine, but acquire it through dietary intake rather than producing it themselves.[38] The absence of the (*S*)-lysine/DAP biosynthetic pathway in mammals means that inhibitors of this pathway would not be expected to have mammalian toxicity.[38]

Gene knockout experiments suggest that *M. tuberculosis* is inefficient at acquiring (*S*)-lysine, because *M. tuberculosis* H37Rv (*S*)-lysine auxotrophs required media supplemented with 25-fold higher concentration of (*S*)-lysine (1 $mg.mL^{-1}$) than similar *M. smegmatis* auxotrophs.[50] The inability of *M. tuberculosis* to uptake (*S*)-lysine from the environment also seemed evident in research showing that *M. tuberculosis* growth was unaffected by a toxic (*S*)-lysine analogue.[50] The difficulties in (*S*)-lysine uptake observed for *M. tuberculosis* would make it particularly vulnerable to inhibition of the enzymes critical for (*S*)-lysine biosynthesis. Taken together, the DAP/(*S*)-lysine pathway, being both essential for bacterial viability and absent in animals, provides an excellent target for antibiotic design, especially in *M. tuberculosis*.

1.2.3 Regulation of DHDPS and the (*S*)-lysine pathway

DHDPS, as the enzyme catalyzing the first committed step of the (*S*)-lysine/DAP biosynthetic pathway **(Figure 1.3)**, is well positioned to have an important role in regulating metabolic flux. The catalytic activity of DHDPS is inhibited, to varying degrees of sensitivity, by (*S*)-lysine (as detailed in section 1.2.4).[51] Plant enzymes are orders of magnitude more sensitive than enzymes from Gram-negative bacteria, such as *E. coli*.[51] Gram-positive bacteria, such as *Bacillus licheniformis*, have DHDPS that are uninhibited by physiological concentrations of (*S*)-lysine.[52] The regulation of DHDPS activity by (*S*)-lysine, in Gram-negative bacteria, reflects the importance of DHDPS in regulating flux of metabolites through the DAP pathway and is consistent with its role as the first unique step in this pathway.

In *E. coli*, the expression of the *dap*A gene encoding DHDPS is regulated by the levels of DAP present in the cell,[53] in contrast to the other enzymes in the (*S*)-lysine pathway which are regulated by the levels of (*S*)-lysine.[54,55] Gene-knockout experiments with *M. smegmatis* revealed differences in the gene regulation of the (*S*)-lysine pathway in mycobacteria compared with *E. coli*, because DAP auxotrophs could not be isolated without silencing of an additional gene, *lys*A, which codes for DAP decarboxylase, an enzyme that converts DAP

into (*S*)-lysine.[25] The expression of the *lys*A gene is not regulated by levels of (*S*)-lysine in *C. glutamicum*,[56] which is considered to be a close evolutionary relative to mycobacteria.[25] Without regulation of *lys*A gene expression, DAP auxotrophs have insufficient DAP to support cell wall synthesis because the majority of DAP acquired is converted into (*S*)-lysine.[25] The unregulated conversion of DAP to (*S*)-lysine indicates that decreasing the metabolic flux through this pathway, by partial inhibition of DHDPS, may aid in mycobacterial cell death, further validating the enzymes of the (*S*)-lysine/DAP biosynthetic pathway as good targets for antibiotic design in mycobacteria.

1.2.4 Trends for (*S*)-lysine inhibition

The orthologues of DHDPS were first grouped into three classes based on their sensitivity to (*S*)-lysine inhibition by Blickling (*et al.* 1997).[57] Plant DHDPS enzymes are strongly inhibited (IC_{50} between 10 and 50 μM),[57] suggesting their regulatory role in controlling metabolic flux through the (*S*)-lysine biosynthetic pathway.[58] The expression of (*S*)-lysine-insensitive DHDPS in plants results in an accumulation of (*S*)-lysine, providing further evidence for its regulatory role.[59,60]

Bacterial DHDPS enzymes are separated into weakly inhibited (~100 fold less sensitive than plant enzymes) and uninhibited classes of DHDPS, associated with Gram-negative and Gram-positive bacteria, respectively.[57] The regulatory role of the weak inhibition in bacterial DHDPS is unclear.[58,60] Strong selective agents for (*S*)-lysine-overproducing mutants failed to induce expression of modified, less (*S*)-lysine sensitive DHDPS in *E. coli*, unlike similar experiments performed with plants.[60] This suggests that the (*S*)-lysine feedback inhibition of *E. coli* DHDPS is not the main point of regulation for the (*S*)-lysine biosynthetic pathway.[60]

The same (*S*)-lysine binding sites have been identified both in weakly inhibited bacterial enzymes and strongly inhibited plant enzymes, using X-ray crystallography studies.[61,62] However, in plant enzymes, key active-site residues shift upon (*S*)-lysine binding suggesting a possible mechanism for inhibition, whereas the mechanism is still unclear for bacterial enzymes.[51,61] The differences in (*S*)-lysine inhibition have been linked to the differences in quaternary structure between plant and bacterial enzymes, which will be discussed in more detail in section 1.5.4.

Even though the mechanism and regulatory role is unclear for (*S*)-lysine inhibition in bacterial DHDPS, most bacterial DHDPS studied to date fit the trend between Gram-negative and Gram-positive bacteria **(Table 1.1)**. These differences in inhibition between Gram-negative and Gram-positive bacteria may be connected to the differences in their cell walls; either their thickness and/or their composition, using DAP or (*S*)-lysine, respectively (see Figure 1.2, section 1.2.2).

Table 1.1: (*S*)-Lysine inhibition of DHDPS activity for various bacterial DHDPS enzymes.

Gram	*Species*	IC_{50}[a]	Inhibition @ [lysine]		Ref.
+	*Bacillus brevis*	---	None	5 mM	63
+	*Bacillus cereus*	---	10 %	30 mM	64
+	*Bacillus licheniformis*	---	None	10 mM	52
+	*Bacillus megaterium*	---	None	13 mM	65,66
+	*Bacillus sphaericus*	0.6 mM	100 %	10 mM	29
+	*Bacillus stearothermophilus*	---	None	10 mM	67
+	*Bacillus subtilis*	---	None	100 mM	68
+	*Brevibacterium flavum*	---	None	---	69
+	*Brevibacterium linens*	---	20 %	20 mM	69
+	*Brevibacterium lactofermentum*	---	None	10 mM	70
+	*C. glutamicum*	660 mM	None	---	71,72
-	*E. coli*	1 mM	50 %	1 mM	73
+	*Lactobacillus plantarum*	---	None	10 mM	74
-	*Methylophilus methylotrophus*	1 mM	90 %	50 mM	75
-	*Methanobacterium thermoautrophicum*	0.25 mM	80 %	20 mM	76
-	*Neisseria meningitidis*	0.053 mM	50 %	0.053 mM	77
-	*Pseudomonas acidovorans L6*	0.20 mM	50 %	0.20 mM	78
-	*Pseudomonas putida T1*	0.65 mM	50 %	0.65 mM	78
-	*Sinorhizobium meliloti*	0.7 mM	50 %	0.7 mM	79
+	methicillin-resistant *Staphylococcus aureus*	225 mM	None	10 mM	80,81
+	*Streptococcus faecalis*	---	None	16 mM	82
+	*Streptomyces clavuligerus*	---	None	10 mM	83
+	*Thermoanaerobacter tengcongensis*	---	None	20 mM	84
-	*Thermotoga maritima*	---	None	10 mM	85

[a] The inhibitor concentration which reduces enzyme activity by 50 %.

1.3 The catalytic activity/enzymatic function of DHDPS

A common approach to rationally designing enzyme inhibitors is through identifying and understanding the chemistry and structure of their active site.[86] Putative inhibitors often have

structures based on the substrates, products or intermediates of the reaction.[86] DHDPS catalyzes an aldol condensation, that is the formation of a β-hydroxy group from two carbonyl-containing compounds, in this case, pyruvate and (*S*)-ASA **(Figure 1.3)**. Research focused on designing an inhibitor for DHDPS has resulted in extensive work elucidating the details of its catalytic mechanism.

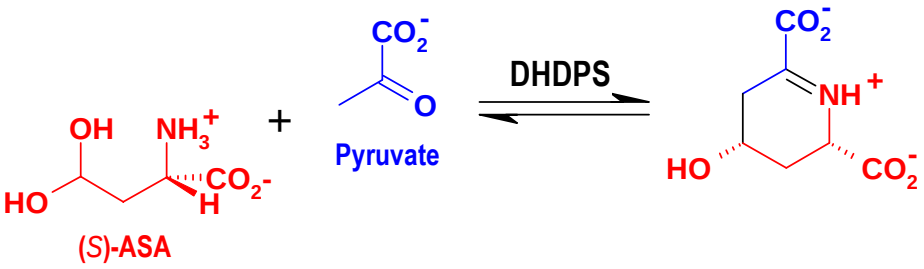


Figure 1.3: The overall reaction catalyzed by dihydrodipicolinate synthase (DHDPS).

1.3.1 Kinetics of catalysis

Numerous kinetic studies of *E. coli* DHDPS suggest that catalysis proceeds through a ping-pong mechanism.[58,73,87-89] This mechanism describes two substrates binding in a specific order, with a release of a product before the binding of the second substrate.[90] Thus the enzyme (E) never contains two unreacted substrates, but the first substrate, in this case pyruvate (PYR), reacts to form an enzyme-substrate intermediate (F), with release of H_2O and it is this intermediate (F) which binds to and then reacts with (*S*)-ASA **(Figure 1.4)**.[89,91]

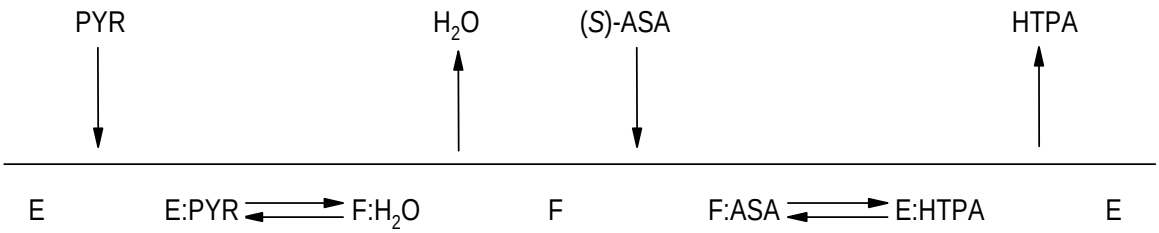


Figure 1.4: **A schematic diagram of the kinetic mechanism of DHDPS, with substrates, pyruvate (PYR) and (*S*)-ASA, binding in a specific order.[89] (4*S*)-4-Hydroxy-2,3,4,5-tetrahydro-(2*S*)-dipicolinic acid (HTPA) is currently thought to be the condensation product released from DHDPS.[57]**

The ping-pong mechanism has been assigned to DHDPS from various plant species,[92-94] *C. glutamicum*,[71] *E. coli*,[58,73,87,88] methicillin-resistant *Staphylococcus aureus* (*MRSA*),[80] and *Thermotoga maritima*,[85] based on the characteristic parallel lines shown by the double reciprocal plots of initial velocity data.[91] The existence of an enzyme-pyruvate intermediate has been demonstrated by mass spectrometry[95] and crystallography,[73,80,96,97] and is consistent with the inhibition pattern shown by some substrate and product analogues, which reflected the existence of two different enzyme forms (E & F) as shown in **Figure 1.4**.[89]

The relationship between substrate concentration and the rate of an enzyme-catalyzed reaction is reflected in the Michaelis constant, K_M, which is determined by fitting initial velocity data to the appropriate mechanistic model.[91] There is variation in reported K_M values for DHDPS from different sources **(Table 1.2)**. Importantly, as noted in **Table 1.2**, several different methods have been used for collecting kinetic data for DHDPS (described in detail in chapter two, section 2.2) and two of those methods, referred to as the *o*-aminobenzaldehyde assay and the imidazole assay, have inherent lag times. Thus it is questionable whether the rate being measured is actually the initial-rate of DHDPS.[98] The coupled assay has no lag time, is the current method of preference,[98] and has repeatedly determined K_M values for *E. coli* DHDPS of similar magnitudes: 0.2 mM for pyruvate and 0.1 mM for (*S*)-ASA **(Table 1.2)**.

The superiority of the coupled assay relative to *o*-aminobenzaldehyde assay is emphasized in the 10-fold difference in the K_M for (*S*)-ASA **(**0.63 as opposed to 6.2 mM, **Table 1.2)** determined for DHDPS from *C. glutamicum*.[71] *E. coli* DHDPS has relatively smaller values of K_M compared with *C. glutamicum* DHDPS **(**~0.2 vs. 0.63 mM for pyruvate, **Table 1.2)**, showing that lower concentrations of substrates are required to saturate the *E. coli* enzyme.

Table 1.2 Kinetic constants for wild-type DHDPS from various sources.

Organism	K_M pyruvate (mM)	K_M (S)-ASA (mM)	Assay method	Ref.
Plants				
Pisum sativum (Pea)	1.7	0.4	*o*-aminobenzaldehyde	92
Spiniacia oleracea (Spinach)	-	1.4	*o*-aminobenzaldehyde	99
Triticum aestivum (Wheat)	11.8	0.8	*o*-aminobenzaldehyde	94
Triticum aestivum (Wheat)	-	1	*o*-aminobenzaldehyde	100
Zea mays (Maize)	2.1	0.6	*o*-aminobenzaldehyde	93
Bacteria				
Bacillus licheniformis	2.6	5.3	*o*-aminobenzaldehyde	101
Bacillus megaterium	0.5	0.46	imidazole	101
Bacillus sphaericus	9	5.1	*o*-aminobenzaldehyde	101
Bacillus subtilis	1.07	3.13	*o*-aminobenzaldehyde	68
Brevibacterium lactofermentum	4.4	0.58	*o*-aminobenzaldehyde	102
C. glutamicum	-	6.2	*o*-aminobenzaldehyde	72
C. glutamicum	0.32 (±0.01)	0.63 (±0.04)	coupled	71
E. coli	0.25	0.13	coupled	23
E. coli	0.57	0.55	imidazole	73
E. coli	0.17	0.07	coupled	89
E. coli	0.25 (±0.03)	0.11 (±0.01)	coupled	103
E. coli	0.26 (±0.03)	0.11 (±0.01)	coupled	87
E. coli	0.14 (±0.03)	0.12 (±0.01)	coupled	88
E. coli	0.16 (±0.03)	0.13 (±0.02)	coupled	2
Methanobacterium thermoautrophicum	0.83	22.2	*o*-aminobenzaldehyde	76
Neisseria meningitidis	0.50 (±0.03)	0.052 (±0.03)	coupled	77
Sinorhizobium meliloti	0.27 (±0.02)	0.13 (±0.02)	imidazole	79
S. aureus (MRSA)	0.11 (±0.01)	0.22 (±0.02)	coupled	81
S. aureus (MRSA)	0.12 (±0.01)	0.33 (±0.03)	coupled	80
Streptomyces clavuligerus	1	0.3	imidazole	83
Thermoanaerobacter tengcongensis[a]	0.85 (±0.05)	0.38 (±0.05)	coupled[a]	84
T. maritima	0.05 (±0.01)	0.16 (±0.01)	coupled	85

[a] Coupled assays were performed at 60 °C rather than 30 °C, nearer the *in vivo* temperatures for the thermophilic bacteria *Thermoanaerobacter tengcongensis*. An enzyme's K_M has been shown to change with changing temperature, therefore the elevated K_M in comparison to others determined by the coupled assay may reflect mainly a temperature effect.[84]

1.3.2 Binding of pyruvate

The reaction catalyzed by DHDPS begins with the ε-amino group of a lysine (K161 in *E. coli* DHDPS) attacking the partially positive carbon of the carbonyl group on pyruvate **(Step 1, Figure 1.5)**.[61] The resulting dehydration to form an imine, or Schiff base **(Step 2, Figure 1.5),** has been thoroughly demonstrated in several studies.[73,95,104] In these studies, sodium

borohydride inactivates DHDPS in the presence of pyruvate, by acting as a strong reducing agent that converts the reactive Schiff base into a stable amine adduct.[104,105] The inactivated protein with a reduced imine adduct at its active-site lysine has been shown by both crystallography[73] and mass spectrometry.[95]

A protonated Schiff base, acting as an "electron sink", is one of the key characteristics of the DHDPS-like superfamily of enzymes.[106] The enzymes of the superfamily, also called the *N*-acetylneuraminate lyase (NAL) family, all contain lysine in their active site for Schiff base formation with substrates containing an α-keto acid group, such as pyruvate.[107] In comparison to DHDPS, the enzyme NAL catalyzes the reverse type of reaction, aldol cleavage rather than condensation, resulting in pyruvate and *N*-acetyl-D-mannosamine.[107]

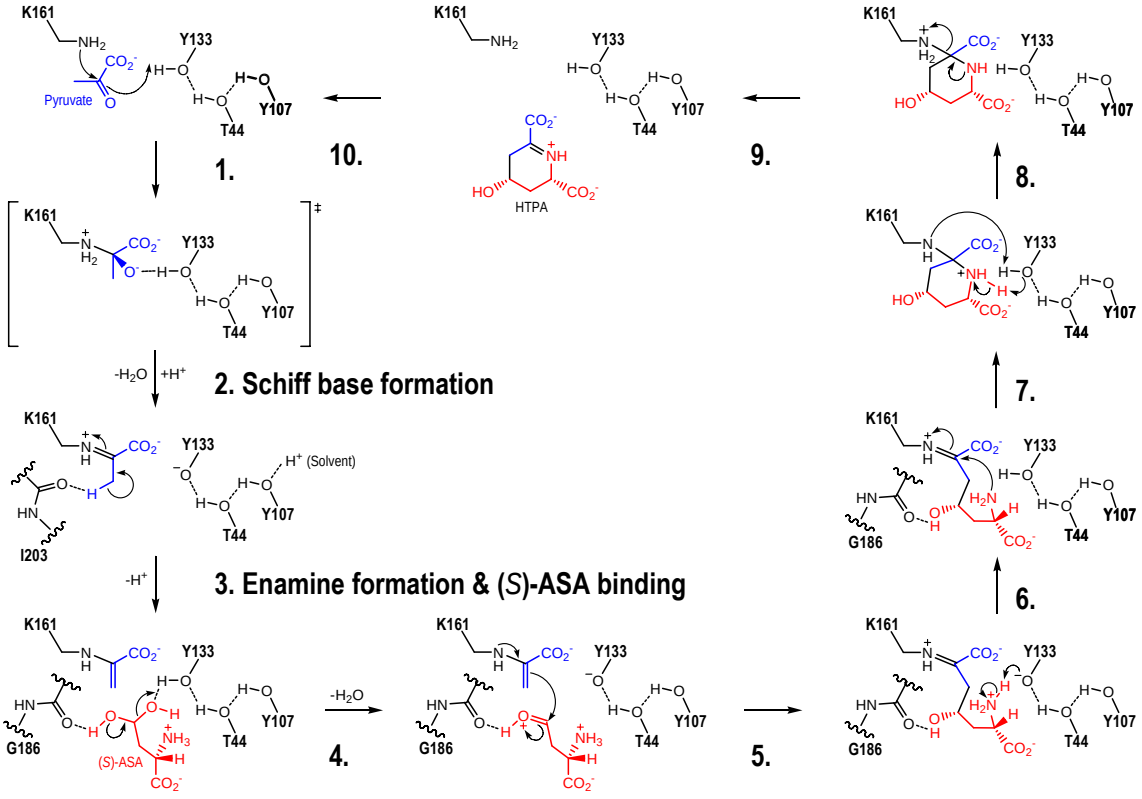


Figure 1.5: The currently accepted mechanism for *E. coli* DHDPS.[58]

Interestingly, NAL has shown low levels of DHDPS activity,[108] and crystallization experiments of DHDPS and NAL with pyruvate and analogues have helped identify other residues important for Schiff base formation/stabilization.[107,109,110] The position of the

substrate's carboxyl group indicated hydrogen bond formation with backbone amides from two nearby amino acids (threonine 44, threonine 45 in *E. coli* DHDPS, and serine 47, threonine 48 in *E. coli* NAL) **(Figure 1.6)**.[57,107,109] These amino acids are part of glycine-x-x-glycine motif highly conserved across the NAL superfamily, and it has been suggested that these glycines have an important structural role in maintaining the orientation of the backbone amides.[110-112]

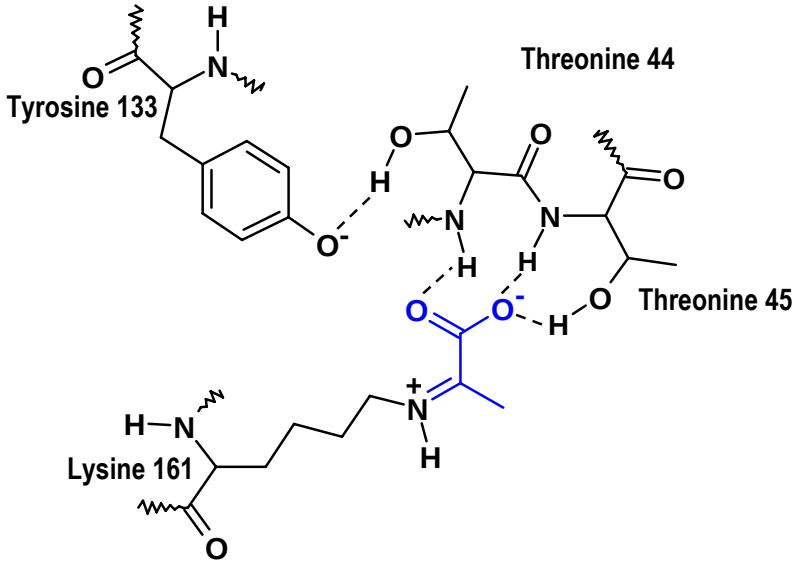


Figure 1.6: **Illustration of interactions between *E. coli* DHDPS and pyruvate (shown as an imine in blue), after Schiff base formation occurs (outlined in Step 2, Figure 1.5). Pyruvate reacts to form the Schiff base via the tetrahedral intermediate (outlined in Step 1, Figure 1.5).**

In addition, crystallography studies with members of the NAL family have identified two more interactions that stabilize bound substrate, involving the hydroxyl containing side chain of the previously mentioned threonine (T45 in *E. coli* DHDPS) and a conserved tyrosine (Y133 in *E. coli* DHDPS).[57,107,109,111] The hydrogen bond geometry for Y133 in *E. coli* DHDPS is not ideal for stabilizing the bound pyruvate; however, Y133 has been proposed to assist the reaction by acting as both a proton donor and acceptor during Schiff base formation **(Figure 1.5)**.[57,110] This role was supported in DHDPS by the observation of a hydrogen bonding network, which was postulated to provide a "back door" through which protons can be shuttled in and out of the active site.[57] Y133 was observed to be hydrogen bonded *via* the hydroxyl group of T44 to a second conserved tyrosine (Y107, in *E. coli* DHDPS), which reaches into the active site from the adjacent subunit.[57] The essential nature of this catalytic

triad has been demonstrated by using site-directed mutagenesis to produce three mutant *E. coli* DHDPS enzymes, DHDPS-Y133F, DHDPS-T44V, and DHDPS-Y107F, all of which had substantially reduced activity.[87] However, only DHDPS-Y133F showed a major change in its K_M value for pyruvate,[87] which is consistent with the hydroxyl group of this residue having a specific role in binding pyruvate, as reflected in the enzyme-bound intermediate drawn in **Figure 1.5**.

1.3.3 Formation of the reactive species

Following Schiff base formation, the imine is converted to an enamine in order to create the necessary reactive species **(Step 3, Figure 1.5)**.[106] This requires proton removal from the methyl group leading to double bond migration, but it is not clear which residue(s) facilitates the proton abstraction.[96] The main-chain oxygen of isoleucine (I203 in *E. coli* DHDPS) has been suggested to assist due to its orientation and proximity, 3.4 Å from the methyl carbon (C3) of the imine, in crystal structures.[57,61,96] Another candidate for mediating proton removal is the member of the catalytic triad Y133; however, this is more than 4 Å from the methyl carbon (C3) of the imine in the crystal structure of *E. coli* DHDPS.[113]

Significant polypeptide backbone strain is found between I203 and the adjacent serine (S204 in *E. coli* DHDPS) in all structurally and biochemically characterized DHDPS structures, suggesting a role in catalysis.[96] The distortion in planarity across this peptide bond has an estimated energy cost of ~3.5 kcal.mol^{-1} and is possibly caused by hydrogen bonds between S204 and an aspartate (D188 in *E. coli* DHDPS) adjacent to the aspartate (D187 in *E. coli* DHDPS) postulated to be involved (*S*)-ASA binding.[96] A similar feature is not observed in other members of the DHDPS-like superfamily indicating it is associated with the binding of (*S*)-ASA.[96,107] The rearrangement of D187 required for its interaction with (*S*)-ASA is postulated to have a cascading effect, causing the rearrangement of D188, S204, I203 and the waters bound to them, and thus increasing the basic nature the main chain carbonyl of I203 and resulting in proton abstraction.[96] Site-directed mutagenesis studies have been proposed to provide some evidence for this mechanism, although none have been reported to date.

1.3.4 Binding of (*S*)-ASA

Several residues suggested to bind (*S*)-ASA are also implicated in improving the substrate's electrophilicity,[38] since in aqueous solution the majority of (*S*)-ASA is known to exist as a

hydrate rather than an aldehyde.[114] Initially, the hydrated form of (*S*)-ASA was postulated to be stabilized by T45 and an asparagine (N248 in *E. coli* DHDPS); however, the latter residue is replaced by glycine in Gram-positive bacteria.[57,61] Further X-ray crystallographic studies of DHDPS with pyruvate and (*S*)-ASA analogues suggested two residues proximate to the active-site lysine residue that could convert (*S*)-ASA into a more reactive aldehyde form.[57] A tyrosine (Y133 in *E. coli* DHDPS) has its side chain in the correct orientation to donate a proton to one hydroxyl group, and a glycine (G186 in *E. coli* DHDPS) has a backbone oxygen in the correct orientation to hydrogen bond with the other hydroxyl group, which cumulatively results in the loss of water and the re-formation of the oxygen double bond of the aldehyde **(Step 4, Figure 1.5)**.[38]

Analysis of X-ray crystallography data has identified additional residues that may interact with (*S*)-ASA, including the active-site lysine K161, which is within hydrogen bonding distance to stabilize it in its aldehyde form.[57] Also, D187 (referred to in section 1.3.3) was suggested to hydrogen bond to the ammonium group of (*S*)-ASA.[107] A highly conserved arginine (R138 *E. coli* DHDPS numbering) has the proximity and orientation to coordinate the carboxyl group.[57,115] Interestingly, a mutation to introduce an arginine at the equivalent position to R138 in *E. coli* NAL resulted in increased DHDPS activity. However, this mutation only marginally affected the K_M for (*S*)-ASA, suggesting that R138 may have another unknown role in catalysis.[108,115] The essential nature of R138 was demonstrated by using site-directed mutagenesis to produce two mutant *E. coli* DHDPS enzymes, DHDPS-R138A, and DHDPS-R138H, both of which had reduced activity, similar to the catalytic triad mutants.[115] Values determined for K_M were entirely consistent with R138 being involved with binding of (*S*)-ASA, but not pyruvate. Structural studies suggested R138 stabilizes the catalytic triad; therefore an additional role for R138 was proposed, of modulating the proton-relay during catalysis.[115]

1.3.5 Formation of product

Once both substrates are bound to the enzyme and have been converted into more reactive forms, the nucleophilic enamine intermediate attacks the electrophilic protonated aldehyde **(Step 5, Figure 1.5)**.[38] This is followed by cyclization **(Step 6, Figure 1.5)** and transimination **(Step 7-9, Figure 1.5)**, resulting in the formation of (4*S*)-4-hydroxy-2,3,4,5-tetrahydro-(2*S*)-dipicolinic acid (HTPA).[61] HTPA is a precursor to dihydrodipicolinate (DHDP), the substrate

for the next enzyme in the pathway, DHDPR. It was previously assumed that DHDP was the product of DHDPS, but NMR experiments suggest that HTPA is the final product and non-catalyzed breakdown results in the formation of DHDP.[57]

The ambiguity in the exact nature of the product of DHDPS was postulated to have hindered the development of inhibitors structurally based on the enzyme's substrates and products. The hypothesis that HTPA, not DHDP, was the product of DHDPS sparked the development of a suite of inhibitors based on HTPA.[116-118] Like previous studies of inhibitors based on pyruvate,[39,95,119] (*S*)-ASA,[98,120] and DHDP,[98,121] none of the compounds tested showed potent inhibition.[116-118] Of the HTPA-based inhibitors, the molecule that was the most similar to HTPA showed no inhibition at all.[116] It seems that the active site of DHDPS is highly specific and therefore substrate or product analogues are unlikely to be potent inhibitors.[116,122]

1.4 Targeting quaternary structure for drug design

A new approach to designing anti-microbial agents is being developed which utilizes our growing understanding of the structural underpinnings of enzyme activity by targeting the protein-protein interactions involved in quaternary structure rather than the active site.[123] The association of subunits to form a specific quaternary structure is critical for catalytic activity in many enzymes,[124,125] as it is often required to form the active site; either directly at the protein-protein interface, or indirectly through conformational changes induced by quaternary structure.[124] Dissociation and association into quaternary structure acts as a regulator mechanism in ~15% of oligomeric proteins.[124,126] This potential for regulation and other roles for quaternary structure will be discussed more extensively in chapter three. The following sections will consider the potential for disrupting quaternary structure as an approach to antibacterial design.

1.4.1 Protein-protein interactions as drug targets

Protein-protein interactions have key roles in most biological processes and thus they are an increasingly recognized target for drug design.[127-129] These interactions vary from temporary associations, such as those involved in cell signalling, to the permanent associations, like those of multi-subunit enzymes, where subunit association can be viewed as a continuation of

protein folding.[130,131] Since many disease states, such as cancer, are closely linked to cell signalling, targeting protein-protein interactions has a huge therapeutic potential.[129,132]

Discovering small-molecule drugs that disrupt protein-protein interactions is an enormous challenge since in principle these small organic compounds have to mimic and block the interactions of much larger protein surfaces.[132] Multiple factors contribute to the apparent resistance of protein-protein interactions as pharmaceutical targets including the typical flatness of the interface, the lack of naturally occurring small-molecule starting points for design, the character of existing small-molecule libraries and difficulties in confirming and characterizing small-molecule-protein interactions.[127,129]

Considering these numerous difficulties, the advances in this field are striking, including clinically approved small-molecule drugs for preventing platelet aggregation,[128,129] and anticancer agents which show robust anti-tumor activity and are currently in phase I/II clinical trials.[132,133] The field of drug design based on targeting protein-protein interfaces is still in its infancy,[123,128] but there have been importance advances in understanding which are of particular interest with regards to quaternary structure disruptors.

1.4.2 The interfaces of protein-protein interactions

Interactions between protein surfaces usually involves relatively large interfaces with good steric and electrostatic complementarity to each other.[134] However, despite the large size of interfaces, a study in 1995 of a heterodimeric protein complex found mutation of individual interface residues had greatly varied effects on binding affinity.[135] It was demonstrated that only a small set of contact residues, which were subsequently referred to as “hot spots”, were actually important for maintaining binding and these were found to have dimensions comparable to drug-like small molecules.[132,135] Thus, in theory, an inhibitor can be designed to complement the “hot spot” in a targeted complex and block protein-protein interactions.

An overview of successful inhibitors of protein-protein interactions reveals further insight into protein interfaces. The overriding importance of side-chain interactions is emphasized by the small molecule inhibitors of integrins, interleukin 2, Z-interacting protein A, B-cell lymphoma 2, and human protein double minute 2 and homologues, as the chemical structures of these inhibitors replicated key side-chain interactions but not the peptide backbone of their

protein partners.[129] Also, the shapes of side chains are not precisely duplicated in inhibitors but rather the general hydrophobic and electrostatic characteristics are matched.[129] These observations suggest there is great freedom for a diversity of chemical scaffolds,[129] and highlight the importance of understanding protein-protein interactions for drug design.

1.4.3 Disrupting quaternary structure as an approach to drug design

The first small molecules designed to inhibit protein-protein interactions were for cell surface receptors called integrins.[129] This research lead to two new anti-platelet drugs, a cyclic peptide, epifibatide, and a synthetic peptide analogue, tirofiban, which are currently used in the treatment of patients with acute coronary syndromes.[128,129,136] Even though it was the functional roles of protein-protein interactions in cell signalling that propagated study, disrupting quaternary structure was also investigated, as some cell signalling protein systems are regulated by oligomerization.[127,129] For example nitric oxide synthase (NOS),[137] and tumour-necrosis factor (TNF),[138] are only functional as dimers and trimers, respectively.

Unexpectedly, inhibitors discovered for NOS were found to inhibit oligomerization by binding not to the protein interface, but rather a pocket distal to the interface, inducing conformational changes that radiate outwards, ultimately distorting the structure of the interface.[129,137] Unfortunately, these inhibitors were found to be ineffective against partially purified dimeric enzyme, thus it is thought that they bind the monomeric form of the enzyme during synthesis.[123,137] The small molecule inhibitor discovered for TNF was shown to bind the trimer and actively promote the dissociation of one subunit, leaving an inactive dimer form.[129,132,138] X-ray crystallography revealed that this inhibitor was located in the cleft between two of the subunits, acting as a wedge, and thus, similar to the inhibitors of NOS, induces conformational changes that propagate outwards, ultimately disrupting oligomerization.[129,138] Interestingly, several other inhibitors of protein-protein interactions have been found to distort the interface in an allosteric manner.[127] This type of inhibitor may be of particular importance to oligomeric proteins, as interfaces are obviously less accessible in permanent complexes than transient complexes.

The majority of inhibitors of protein-protein interaction function as competitive rather than allosteric inhibitors, like those mentioned in section 1.4.2. Numerous investigations have targeted protein-protein interactions at the interface using peptides which act as starting points

for drug-like small molecule design.[131,139,140] For example, the homodimer protease from HIV was found to be inhibited by peptides corresponding to its interface as early as 1991.[141-143] More recently, AUC experiments with alkyl tripeptides indicated that these optimized inhibitors bind to the interface of the active dimer, promoting dissociation and sequestering HIV protease monomers.[144] This supports the proposal that the dynamic motions of proteins enable small molecules to penetrate protein interfaces, thus the binding of an inhibitor to a protein interface does not necessarily require the complete dissociation of the complex.[132] As of yet there are no commercially available drugs that act by disrupting the quaternary structure of multi-subunit enzymes, but these representative cases highlight both the obstacles and possibilities of this exciting new approach to drug design.

1.5 The structure of DHDPS

The importance of the homotetrameric quaternary structure of DHDPS was explored through site-directed mutagenesis studies with DHDPS from *E. coli* (section 1.5.3), and suggested that molecules designed to disrupt quaternary structure would also inhibit enzyme activity.[1,123,145] Until recently, all structurally and biochemically characterized DHDPS orthologues were tetramers (section 1.5.1). However, newly published work revealed the first native non-tetrameric DHDPS, indicating that the homotetramer quaternary structure is not required for activity in all orthologues (section 1.5.5).[81] This emphasizes the importance of understanding structural details of the enzyme from the pathogen in question when exploring the feasibility of designing protein-protein interaction disruptors as inhibitors for DHDPS.

1.5.1 Structural studies of DHDPS

The solid-state structures of several bacterial DHDPS enzymes have been determined by X-ray crystallography **(Table 1.3)**. Additionally, one plant enzyme, *Nicotiana sylvestris* DHDPS has been solved to a resolution of 2.8 Å **(Figure 1.7B)**.[61] The structures solved for the wild-type *E. coli* DHDPS show it to be a homotetramer, that is, composed of four identical subunits, as shown by crystallographic symmetry **(Figure 1.7A)**. Of the other native structures determined, those shown to have DHDPS activity are all homotetrameric in the crystal, with the notable exception of the recently discovered dimeric *MRSA* DHDPS.[2,81,85]

Table 1.3: Solid-state structures of orthologues of DHDPS, putative DHDPS enzymes and DHDPS mutants determined by X-ray crystallography. (TBP = to be published).

PDB[a] code	Organism	Resolution (Å)	Ligands/mutations	Ref.	Activity assayed
2HMC	*Agrobacterium tumefaciens*	1.90	-	TBP	No
2R8W	*Agrobacterium tumefaciens*	1.80	-	TBP	No
3B4U	*Agrobacterium tumefaciens*	1.20	-	TBP	No
2EHH	*Aquifex aeolicus*	1.90	-	TBP	No
1XL9	*Bacillus anthracis*	2.23	-	146	Yes
1XKY	*Bacillus anthracis*	1.94	-	146	Yes
3E96	*Bacillus clausii*	1.80	-	TBP	No
3CPR	*C. glutamicum*	2.20	-	71	Yes
3BI8	*Clostridium botulinum*	1.96	-	TBP	No
1DHP	*E. coli*	2.30	A207T	62	Yes
1YXC	*E. coli*	1.90	-	58	No
1YXD	*E. coli*	2.00	lysine	58	Yes
2ATS	*E. coli*	1.90	lysine	TBP	No
3C0J	*E. coli*	2.40	hydroxypyruvate	96	Yes
1S5T	*E. coli*	2.30	T44V	87	Yes
1S5V	*E. coli*	2.35	Y107F	87	Yes
1S5W	*E. coli*	2.32	Y133F	87	Yes
2A6L	*E. coli*	2.05	R138H	147	Yes
2A6N	*E. coli*	1.94	R138A	147	Yes
2OJP	*E. coli*	1.70	L197Y	2	Yes
3DU0	*E. coli*	2.00	pyruvate	113	No
2PUR	*E. coli*	1.70	T44S	TBP	No
3DEN	*E. coli*	2.20	Y107W	148	Yes
2RFG	*Hahella chejuensis*	1.50	-	TBP	No
2YXG	*Methanocaldococcus jannaschii*	2.20	-	TBP	No
1XXX	*M. tuberculosis*	2.28	-	149	Yes
3FLU	*Neisseria meningitidis*	2.00	-	77	Yes
-	*Nicotiana sylvestris*[a]	2.79		61	No
-	*Nicotiana sylvestris*[a]	2.80	lysine	61	No
3D0C	*Oceanobacillus iheyensis*	1.90	-	TBP	No
3DZ1	*Rhodopseudomonas palustris*	1.87	-	TBP	No
3EB2	*Rhodopseudomonas palustris*	2.04	-	TBP	No
3G0S	*Salmonella typhimurium*	1.85	-	TBP	No
2VC6	*Sinorhizobium meliloti*	1.95	pyruvate	97,150	Yes
3DI0	*S. aureus (MRSA)*	2.38	-	80	Yes
3DI1	*S. aureus (MRSA)*	2.20	pyruvate	80	Yes
3DAQ	*S. aureus (MRSA)*	1.45	-	81	Yes
1O5K	*T. maritima*	1.80	-	151	Yes
2PCQ	*Thermus thermophilus*	2.10	-	TBP	No

[a] Most structural data is freely availbable from Protein Data Bank (PDB).[152] *Nicotiana sylvestris* DHDPS is not deposited.

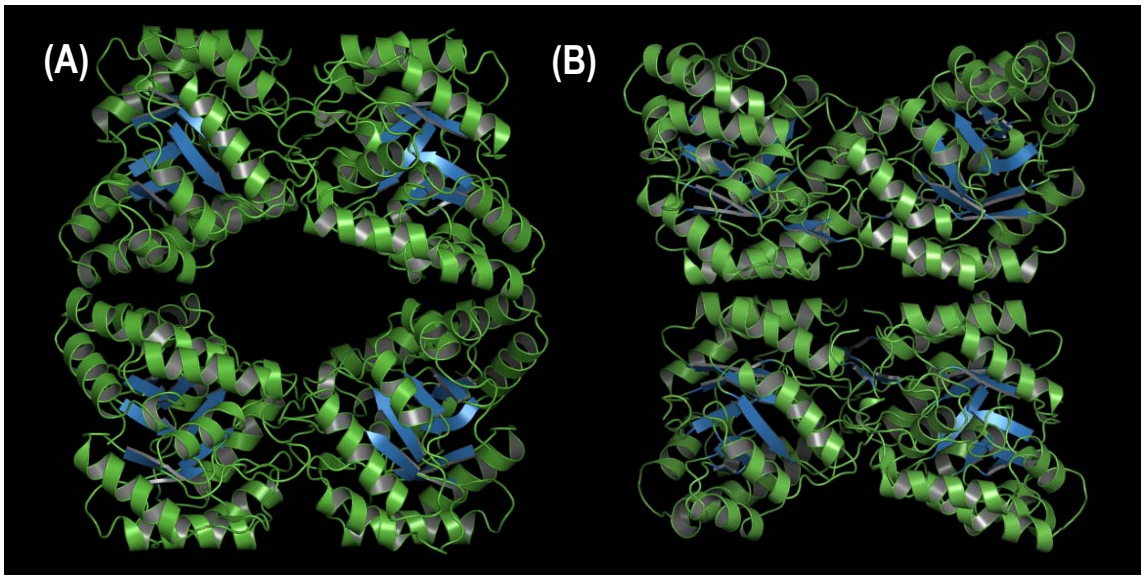


Figure 1.7: The differences in quaternary structure between (A) bacterial *E. coli* DHDPS and (B) plant *N. sylvestris* DHDPS. Strong protein-protein interactions form analogous "tight-dimers", however these associate *via* opposite interfaces into two alternative tetramers.[61] Figures illustrating structural details have been prepared using the program PyMOL.[153]

The putative DHDPS from *Agrobacterium tumefaciens* (PDB entry 2HMC) seems to be a hexamer (generated by crystallographic symmetry) but is likely misclassified as it does not appear to possess all of the key catalytic residues of DHDPS.[96,149] Additionally, the putative DHDPS from *Thermus thermophilus* (PDB entry 2PCQ) may also be misclassified, as there is accumulating evidence that the bacteria utilizes the α-aminoadipate pathway found in fungi to synthesize (*S*)-lysine rather than the DAP pathway common to bacteria.[154-156] These possible misclassifications emphasize the importance of performing functional studies in conjunction with structural determination.

Investigation with biophysical methods, such as gel-filtration liquid chromatography, has found DHDPS to be tetrameric, with two exceptions **(Table 1.4)**. Contrary to other plant enzymes, gel filtration found DHDPS from *Pisum sativum* to be trimeric,[92] but this anomalous result has not been confirmed by other methods. A more quantitative biophysical method, analytical ultracentrifugation (AUC), was used to confirmed the dimeric nature of *MRSA* DHDPS, which has similar activity to the tetrameric *E. coli* DHDPS.[2,145]

Table 1.4: Oligomeric states of DHDPS from various sources.

Organism	Bacterial/Plant	Oligomeric state	Ref.
Bacillus licheniformis	Bacterial	tetramer	101
Bacillus subtilis	Bacterial	tetramer	68
C. glutamicum	Bacterial	tetramer	71
E. coli	Bacterial	tetramer	73,95,104
N sylvestris	Plant (Tobacco)	tetramer	157
Pisum sativum	Plant (Pea)	trimer	92
Sinorhizobium meliloti	Bacterial	tetramer	97
S. aureus (MRSA)	Bacterial	dimer	80,81
Thermoanaerobacter tengcongensis	Bacterial	tetramer	84
T. maritima	Bacterial	tetramer	85
Triticum aestivum	Plant (Wheat)	tetramer	94
Zea mays	Plant (Maize)	tetramer	93

The structural features of DHDPS have been well characterized, and it has been described as a dimer of "tight-dimers" **(Figure 1.7A)**, a quaternary structure which is conserved in most bacterial species studied.[148] Curiously, the plant enzyme from *N. sylvesteris* DHDPS adopts a similar functional unit, composed of two tight-dimers that interact to form a tetramer, although the arrangement of the tight-dimers is completely different **(Figure 1.7B).**[61] The existence of two alternative tetramers with essentially the same dimeric subunit, suggest that the tetrameric form gives a generic structural advantage rather than specifically contributing to the active site formation.[158] This leads to the hypothesis that DHDPS evolved from an ancestral dimeric protein.[61,85]

Interestingly, DHDPS from *T. maritima*, one of the slowest evolving lineages of eubacteria, has been demonstrated unequivocally to be a tetramer,[85] which suggests that bacterial and plant enzymes evolved from a dimeric enzyme that predates *T. maritima*.[85] Sequence alignments between bacterial DHDPS and the archaeal DHDPS from *Methanocaldococcus jannaschii* suggested *M. jannaschii* DHDPS could be dimeric, and thus would provide insight into the protein evolution of DHDPS.[61] Archaea are considered by many to be the most ancient organisms on earth;[159] thus, archaeal enzymes may provide the closest modern relative of ancestral DHDPS. However, the recently deposited structure of a putative DHDPS from *M. jannaschii* (PDB entry 2YXG) shows it to be tetrameric in the crystal **(Table 1.3)**. Biophysical techniques are needed to confirm that this enzyme is tetrameric in solution, as well as functional studies to confirm DHDPS activity.

In order to explore the hypothesis of a dimeric ancestral DHDPS, a discrete tight-dimer from *E. coli* DHDPS was engineered using site-directed mutagenesis (discussed in section 1.5.3).[2]

These investigations led to the hypothesis that an ancient monomeric form of the enzyme existed,[2] and attempts are ongoing to engineer a monomeric variant of *E. coli* DHDPS.[148] The tertiary structure of the monomer between bacterial species is conserved, even though the bacterial sequences are strongly divergent.[61,62] The structure of *M. tuberculosis* DHDPS had been deposited in the PDB at the outset of this work, but had not been published or biochemically characterized (to be discussed in chapter 2). Consequently, the best studied bacterial DHDPS, from *E. coli*, will be discussed in detail in the following sections.

1.5.2 The tertiary structure of *E. coli* DHDPS

The monomeric unit from *E. coli* DHDPS has a molecular mass of 31272 Da, as determined by electrospray mass spectrometry, which is in agreement with the mass predicted from the amino-acid sequence.[87,95] It contains two domains: a $(\beta/\alpha)_8$-barrel N-terminal domain and an α-helical C-terminal domain **(Figure 1.8A)**.[62] The active-site lysine, K161, is found at the C-terminal end of the $(\beta/\alpha)_8$-barrel, in a 30 Å long by 10 Å deep solvent-accessible pocket **(Figure 1.8B)** that is partially blocked by the adjacent, closely associated monomer.[58,62]

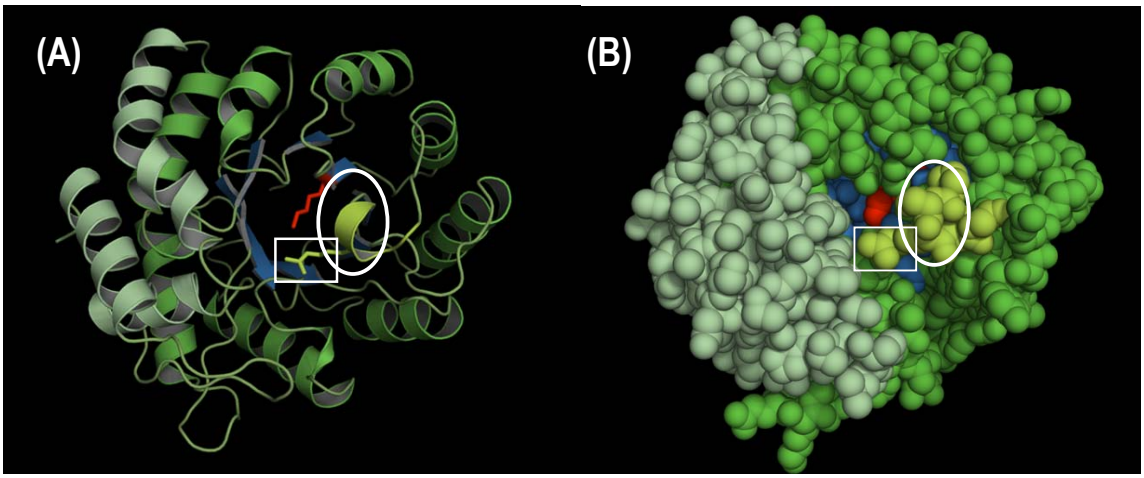


Figure 1.8: The structure of the *E. coli* DHDPS monomeric subunit, as viewed from the C-terminus end of the $(\beta/\alpha)_8$-barrel. (A) The tertiary structure contains the common $(\beta/\alpha)_8$-barrel fold, composed of β-sheets (blue) surrounded by α-helices (green) with an extra α-helix (yellowy green, circled in white) and mainly α-helical region (lighter green). (B) The solvent-accessible pocket containing the active-site residue K161 (red) is more apparent when using a space-filling model. The extra α-helix (circled in white) containing the active-site residue R138 (white square), lies over the middle of the pocket.

The active site is lined with amino acids that orientate substrates, as well as those that directly contribute to the chemistry,[160] as was described in section 1.3.2-4, and most of these active-site residues line the β-strands at the C-terminal end of the $(\beta/\alpha)_8$-barrel.[62] The $(\beta/\alpha)_8$-barrel structure of the N-terminal domain is a common structural fold first identified in triosephosphate isomerase (TIM).[161] DHDPS contains a small departure from this common fold with a short extra stretch of helix (residues 136 to 139) that lies perpendicular to the $(\beta/\alpha)_8$-barrel on the C-terminal side **(Figure 1.9)**.[62] This helix partially blocks the active-site pocket but contains the active-site residue arginine, R138, which is proposed to coordinate the carboxyl group of (*S*)-ASA.[61,115] In contrast, the C-terminal domain, which is composed of three α-helices, has no obvious function in catalysis or the regulatory mechanism.[62]

1.5.3 The role of quaternary structure in *E. coli* DHDPS

The homotetramer of *E. coli* DHDPS is commonly described as dimer of dimers,[57,87] which emphasizes the presence of two different sets of protein-protein interactions **(Figure 1.9)**. The monomers closely associate to form dimers and the consequent tight-dimers interact to form a tetramer with a large solvent-filled central cavity.[62]

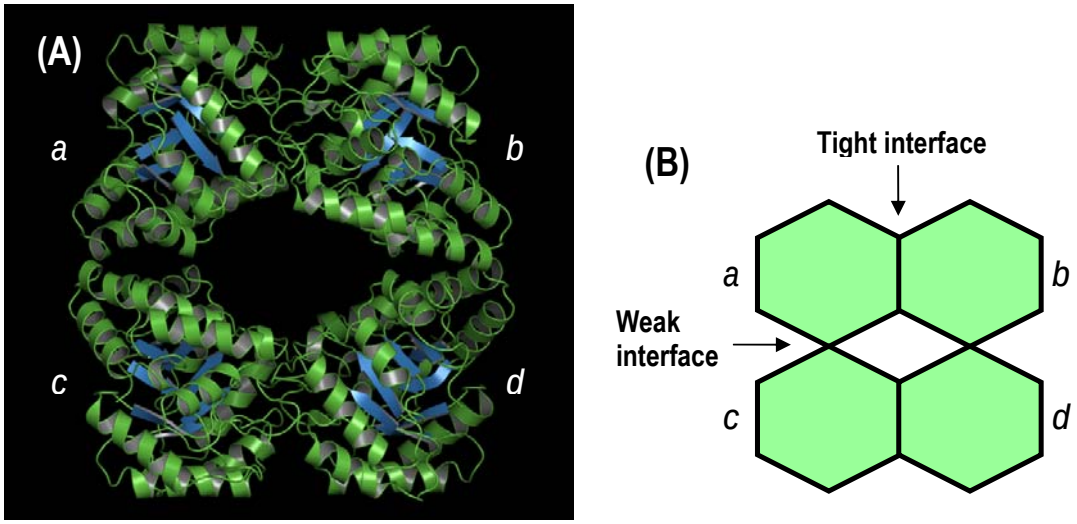


Figure 1.9: **The structure of the *E. coli* DHDPS tetramer. (A) The quaternary structure is a homotetramer with subunits *a* and *b* and subunits *c* and *d* associating strongly to form dimers and the consequent dimers interacting more weakly to form the tetramer. (B) A schematic representation of the two different interfaces referred to as tight and weak, names descriptive of the difference in the number of residues involved in the inter-subunit interaction.**

The tight-dimer interface of the closely associating monomers buries 12.4 % (1400 Å^2) of the total monomeric area,[62] whereas the weak interface between the two dimers buries only 5.8 % (650 Å^2) of the total monomeric area,[61] shown schematically in **Figure 1.9B**. There are four active sites in *E. coli* DHDPS,[104] and all active-site residues are found near the tight interface of the closely associating monomers.[61] The crevice containing the active site is partially blocked by the adjacent monomer, which effectively closes one end of the cleft and completes the catalytic triad with residue Y107 **(Figure 1.10A)**. The catalytic importance of Y107 is reflected in the unusual conformation indicated by it existing in the forbidden region of the Ramachandran plot.[57] The resulting cavity in each monomer is only accessible from the centre of the tetramer.[62] The active-site residue K161 is within the cavity **(Figure 1.10A)**, in close proximity to Y133, which binds pyruvate's carboxyl group.[57,61] Y133 also participates in the catalytic triad with T44 and Y107 (discussed in section 1.3.2), providing a "back door" for shuttling protons into and out of the active site.[57,87] Except for Y107, all active-site residues are contained within the monomeric subunit.[87]

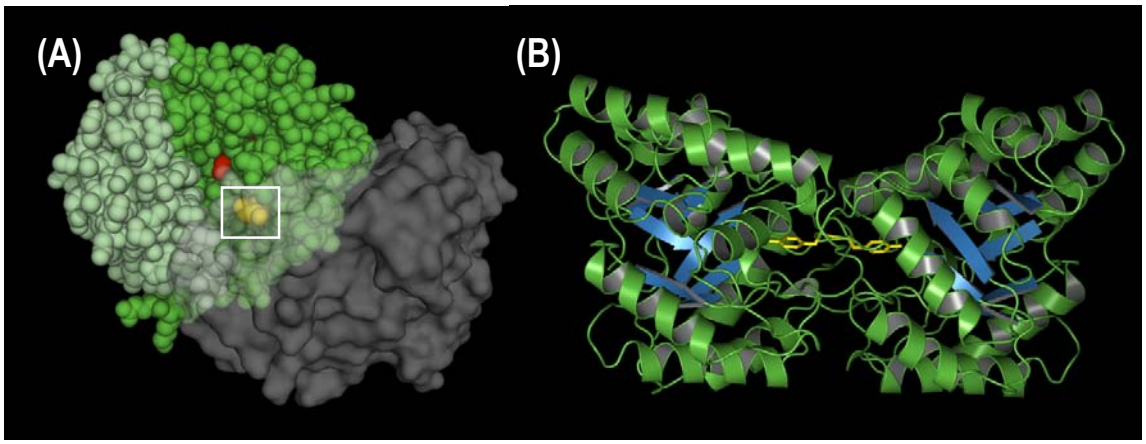


Figure 1.10: The structure of the *E. coli* DHDPS dimeric subunit. (A) Lysine 161 (red) is found in the cleft formed between the two domains (green & light green), which is partially blocked by adjacent monomer (grey shadow) with tyrosine 107 (yellow, in white square), thus creating a cavity only accessible from the centre of the tetramer. (B) Tyrosine 107 (yellow) reaches into the active site from the adjacent monomer in an interdigitating configuration and is the only active-site residue not contained within the monomeric unit.

The interdigitation of the key active-site residue Y107 across the tight-dimer interface **(Figure 1.10B)** provides an explanation for the association of monomers into dimers, but less

clear is the reason for the association of dimers into a homotetrameric quaternary structure.[2] There is no known communication between the dimeric units across the weak interface.[2] In order to explore the importance of the tetrameric structure for *E. coli* DHDPS, dimeric variants, DHDPS-L197Y and DHDPS-L197D were created by site-directed mutagenesis.[2] The activity of these dimeric mutants was substantially less than the wild-type tetramer, even though X-ray crystallography showed the active site to be undisrupted and no significant changes to the tertiary structure.[2] Surprisingly, electron density consistent with a tetrahedral adduct was found at the active-site K161, which was confirmed by mass spectrometry to be α-ketoglutarate, a pyruvate analogue, trapped as a cyclic adduct.[2] The observation of this adduct, believed to have formed *in vivo* since no α-ketoglutarate was added to the enzyme preparation, suggests that the dimeric variants have lower substrate specificity.[2]

The decreased substrate affinity of the dimeric variants was proposed to be related to an increase in dynamic motion, evident with small-angle X-ray diffraction.[2] These dynamic fluctuations mean that the completion of the catalytic triad by Y107 is often disrupted as the monomers reorientate themselves in relation to each other, and this is proposed to cause the reduced binding affinity and specificity for pyruvate observed in the dimeric mutants.[2] Interestingly, the active site of the tetrameric mutant DHDPS-Y107F also had α-ketoglutarate trapped as a cyclic adduct, providing evidence for the importance of Y107 in substrate specificity.[2,87] α-Ketoglutarate is a reversible inhibitor, and incubation at 40 ^{o}C in the presence of pyruvate increased the activity of the dimeric mutants from ~1 % to 10-15 % of wild-type *E. coli* DHDPS and a similar effect was seen with DHDPS-Y107F, presumably due to the removal of the cyclic adduct.[2] The failure of the mutants to gain full wild-type activity hints at the complex role of the catalytic triad in the activity of DHDPS, and provides evidence of the quaternary structure's importance in maintaining the catalytic triad.

1.5.4 The (*S*)-lysine binding site in DHDPS

The binding pocket for the allosteric inhibitor (*S*)-lysine is not limited to the monomeric unit, but found at the tight interface of the closely associating monomers in both *E. coli* and *N. sylvestris* DHDPS.[61] Crystal-soaking experiments show two (*S*)-lysine molecules in the binding pocket within van der Waals contact of each other.[57,58,61] Upon binding of (*S*)-lysine the structure of all four units of *N. sylvestris* DHDPS are significantly changed,[61] but in *E. coli* DHDPS relatively few residues in close proximity to the (*S*)-lysine binding site shift.[58] This

difference results from a dramatic reorientation of the two sets of dimers in relation to each other **(Figure 1.7)** and possibly explains the difference in feedback sensitivity between plant and bacterial DHDPS enzymes.[61] Additionally, there are differing residues coordinating the ϵ-amino group of (*S*)-lysine (H53, H56 in *E. coli* DHDPS vs. W77, H80 in *N. sylvestris* DHDPS), which likely cause a difference in (*S*)-lysine affinity.[51,61]

The mechanism of (*S*)-lysine inhibition in *E. coli* DHDPS is unclear, although kinetic and structural studies support the proposal that it is an allosteric inhibitor.[58] Structural studies of *E. coli* DHDPS with inhibitor bound identified the residues coordinating (*S*)-lysine as Y106, N80, E84, H53, H56, G78, A49, S48; none of which participate in the active site.[57,58] Structural alterations are observed near the allosteric binding site, as various residues move to accommodate the binding of the (*S*)-lysine, the most notable is a tyrosine (Y106), which forms an aromatic stack with a member of the catalytic triad, Y107.[51,58] Y106 shifts towards the carboxyl group of (*S*)-lysine, resulting in altered conformations of both the aromatic stack and Y107, thus likely affecting the catalytic triad, although whether this change is significant enough to be responsible for inhibition is debated.[51,58] More clearly refuted is the proposal that the decreased flexibility of the catalytic residue R138 is a mechanism for inhibition, as subsequent higher resolution structures have shown increased flexibility.[58] While the mechanism for (*S*)-lysine inhibition observed in *E. coli* DHDPS is unknown, the position of the allosteric binding site at the tight-dimer interface further supports the importance of the dimeric subunit.[162]

1.5.5 The quaternary structure of *MRSA* DHDPS

The tetrameric structure of DHDPS is thought to be essential for enzyme activity by reducing the dynamic motion of the dimeric subunit (as discussed in section 1.5.3).[2] Thus it was of particular interest during the course of this research when two separate publication revealed that *MRSA* DHDPS adopts a dimeric quaternary structure in solution, as shown by various biophysical techniques.[80,81] X-ray structural analyses revealed that *MRSA* DHDPS has significantly increased contacts and greater buried surface area at the tight-dimer interface compared to *E. coli* DHDPS.[81] The interdigitating Y109 (equivalent to Y107 in *E. coli* DHDPS), while maintaining the catalytic triad, was found to adopt an altered confirmation in comparison to *E. coli* DHDPS, making a closer aromatic stacking interaction, presumably decreasing its flexibility.[81] These differences suggest an alternative mechanism to overcome

the effect of dynamic motion hypothesized to decrease enzyme activity in the dimeric *E. coli* DHDPS mutants,[81] and indicates the complexity of the interplay between quaternary structure, enzyme function and protein dynamics, which are an increasingly recognized component in enzyme catalysis.[162]

1.6 Summary

M. tuberculosis causes more death than any other bacteria and increasingly with multi-drug resistant strains. Accordingly, there is an urgent need to develop new anti-tuberculosis drugs and characterize novel drug targets, such as DHDPS. While DHDPS does not appear to contribute to variation in virulence or drug susceptibility between strains of *M. tuberculosis*, it is essential for bacterial growth, validating it as a drug target. DHDPS catalyzes a key step in the metabolic pathway yielding the essential amino-acid (*S*)-lysine, and DAP, which forms critical DAP-DAP linkages in the mycobacterial cell wall. DAP auxotrophs of *M. smegmatis*, either underwent cell death in the absence of DAP, or exhibited altered cell wall structures, which rendered them susceptible to antibiotics. In addition, *M. tuberculosis* has a demonstrated inability to uptake environmental (*S*)-lysine, suggesting the bacteria would be particularly vulnerable to inhibitors of DHDPS. The lack of regulation observed in mycobacteria in the conversion of DAP to (*S*)-lysine suggest thats even partial inhibitors of DHDPS may be effective anti-tuberculosis agents, and strengthens the need to fully investigate DHDPS as a drug target.

E. coli DHDPS is the best characterized DHDPS enzyme and through its extensive study an in-depth understanding of the catalytic mechanism is being developed, which has informed the design of inhibitors targeted at the active site. However, it would seem the active site is highly specific, as none of these showed potent inhibition. Therefore, an alternative approach to drug design is considered, targeting protein-protein interactions rather than the active site. Designing inhibitors of protein-protein interaction is more challenging but possible, as indicated by the discovery of "hot spots"; that is, only a small subset of the residues involved in protein-protein interactions are crucial for the association of complexes. Thus, if oligomerization is critical for activity, an inhibitor can be designed to complement the "hot spots" and block subunit association. This thesis will explore the feasibility of targeting the weak interface of *M. tuberculosis* DHDPS as an approach to drug design and consequently, broaden our understanding of the relationship between structure and function in DHDPS.

1.7 References

1. Hutton, C. A., Perugini, M. A. & Gerrard, J. A. Inhibition of lysine biosynthesis: an evolving antibiotic strategy. *Molecular biosystems* **3**, 458-65 (2007).

2. Griffin, M. D., Dobson, R. C., Pearce, F. G., Antonio, L., Whitten, A. E., Liew, C. K., Mackay, J. P., Trewhella, J., Jameson, G. B., Perugini, M. A. & Gerrard, J. A. Evolution of quaternary structure in a homotetrameric enzyme. *Journal of molecular biology* **380**, 691-703 (2008).

3. Laughon, B. E. New tuberculosis drugs in development. *Current topics in medicinal chemistry* **7**, 463-73 (2007).

4. Bloom, B. R. & Murray, C. J. Tuberculosis: commentary on a reemergent killer. *Science* **257**, 1055-64 (1992).

5. Dolin, P. J., Raviglione, M. C. & Kochi, A. Global tuberculosis incidence and mortality during 1990-2000. *Bulletin of the World Health Organization* **72**, 213-20 (1994).

6. Global tuberculosis control - surveillance, planning, financing. in *World Health Organization report* (World Health Organization, Geneva, 2007).

7. Ryan, F. *The forgotten plague: how the battle against tuberculosis was won--and lost* (Little Brown, Boston, 1993).

8. Gandhi, N. R., Moll, A., Sturm, A. W., Pawinski, R., Govender, T., Lalloo, U., Zeller, K., Andrews, J. & Friedland, G. Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area of South Africa. *Lancet* **368**, 1575-80 (2006).

9. Aziz, M. A., Wright, A., Laszlo, A., De Muynck, A., Portaels, F., Van Deun, A., Wells, C., Nunn, P., Blanc, L. & Raviglione, M. Epidemiology of antituberculosis drug resistance (the global project on anti-tuberculosis drug resistance surveillance): an updated analysis. *Lancet* **368**, 2142-54 (2006).

10. Agger, E. M. & Andersen, P. A novel TB vaccine; towards a strategy based on our understanding of BCG failure. *Vaccine* **21**, 7-14 (2002).

11. Fine, P. E. Variation in protection by BCG: implications of and for heterologous immunity. *Lancet* **346**, 1339-45 (1995).

12. Orme, I. M. Tuberculosis vaccines: current progress. *Drugs* **65**, 2437-44 (2005).

13. Orme, I. M. & Belisle, J. T. TB vaccine development: after the flood. *Trends in microbiology* **7**, 394-5 (1999).

14. Cole, S. T., Brosch, R., Parkhill, J., Garnier, T., Churcher, C., Harris, D., Gordon, S. V., Eiglmeier, K., Gas, S., Barry, C. E., 3rd, Tekaia, F., Badcock, K., Basham, D., Brown, D., Chillingworth, T., Connor, R., Davies, R., Devlin, K., Feltwell, T., Gentles, S., Hamlin, N., Holroyd, S., Hornsby, T., Jagels, K., Krogh, A., McLean, J., Moule, S., Murphy, L., Oliver, K., Osborne, J., Quail, M. A., Rajandream, M. A., Rogers, J., Rutter, S., Seeger, K., Skelton, J., Squares, R., Squares, S., Sulston, J. E., Taylor, K., Whitehead, S. & Barrell, B. G. Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome sequence. *Nature* **393**, 537-44 (1998).

15. Camus, J. C., Pryor, M. J., Medigue, C. & Cole, S. T. Re-annotation of the genome sequence of *Mycobacterium tuberculosis* H37Rv. *Microbiology* **148**, 2967-73 (2002).

16. Zheng, H., Lu, L., Wang, B., Pu, S., Zhang, X., Zhu, G., Shi, W., Zhang, L., Wang, H., Wang, S., Zhao, G. & Zhang, Y. Genetic basis of virulence attenuation revealed by comparative genomic analysis of *Mycobacterium tuberculosis* strain H37Ra versus H37Rv. *PLoS one* **3**, e2375 (2008).

17. Fleischmann, R. D., Alland, D., Eisen, J. A., Carpenter, L., White, O., Peterson, J., DeBoy, R., Dodson, R., Gwinn, M., Haft, D., Hickey, E., Kolonay, J. F., Nelson, W. C., Umayam, L. A., Ermolaeva, M., Salzberg, S. L., Delcher, A., Utterback, T., Weidman, J., Khouri, H., Gill, J., Mikula, A., Bishai, W., Jacobs Jr, W. R., Jr., Venter, J. C. & Fraser, C. M. Whole-genome comparison of *Mycobacterium tuberculosis* clinical and laboratory strains. *Journal of bacteriology* **184**, 5479-90 (2002).

18. Cubillos-Ruiz, A., Morales, J. & Zambrano, M. M. Analysis of the genetic variation in *Mycobacterium tuberculosis* strains by multiple genome alignments. *BMC research notes* **1**, 110 (2008).

19. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research* **22**, 4673-80 (1994).

20. Vissa, V. D. & Brennan, P. J. The genome of *Mycobacterium leprae*: a minimal mycobacterial gene set. *Genome biology* **2**, reviews1023.1-.8 (2001).

21. Lamichhane, G., Zignol, M., Blades, N. J., Geiman, D. E., Dougherty, A., Grosset, J., Broman, K. W. & Bishai, W. R. A postgenomic method for predicting essential genes at subsaturation levels of mutagenesis: application to *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 7213-8 (2003).

22. Sassetti, C. M., Boyd, D. H. & Rubin, E. J. Genes required for mycobacterial growth defined by high density mutagenesis. *Molecular microbiology* **48**, 77-84 (2003).

23. Yugari, Y. & Gilvarg, C. The condensation step in diaminopimelate synthesis. *The Journal of biological chemistry* **240**, 4710-6 (1965).

24. Viola, R. E. The central enzymes of the aspartate family of amino acid biosynthesis. *Accounts of chemical research* **34**, 339-49 (2001).

25. Pavelka, M. S., Jr. & Jacobs, W. R., Jr. Biosynthesis of diaminopimelate, the precursor of lysine and a component of peptidoglycan, is an essential function of *Mycobacterium smegmatis*. *Journal of bacteriology* **178**, 6496-507 (1996).

26. Born, T. L. & Blanchard, J. S. Structure/function studies on enzymes in the diaminopimelate pathway of bacterial cell wall biosynthesis. *Current opinion in chemical biology* **3**, 607-13 (1999).

27. McCoy, A. J., Adams, N. E., Hudson, A. O., Gilvarg, C., Leustek, T. & Maurelli, A. T. L,L-diaminopimelate aminotransferase, a trans-kingdom enzyme shared by *Chlamydia* and plants for synthesis of diaminopimelate/lysine. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 17909-14 (2006).

28. Hudson, A. O., Singh, B. K., Leustek, T. & Gilvarg, C. An L,L-diaminopimelate aminotransferase defines a novel variant of the lysine biosynthesis pathway in plants. *Plant physiology* **140**, 292-301 (2006).

29. Bartlett, A. T. M. & White, P. J. Regulation of the enzymes of lysine biosynthesis in *Bacillus sphaericus* NCTC 9602 during vegetative growth. *Journal of general microbiology* **132**, 3169-77 (1986).

30. Paiva, A. M., Vanderwall, D. E., Blanchard, J. S., Kozarich, J. W., Williamson, J. M. & Kelly, T. M. Inhibitors of dihydrodipicolinate reductase, a key enzyme of the diaminopimelate pathway of *Mycobacterium tuberculosis*. *Biochimica et biophysica acta* **1545**, 67-77 (2001).

31. Hudson, A. O., Gilvarg, C. & Leustek, T. Biochemical and phylogenetic characterization of a novel diaminopimelate biosynthesis pathway in prokaryotes identifies a diverged form of LL-diaminopimelate aminotransferase. *Journal of bacteriology* **190**, 3256-63 (2008).

32. Schrumpf, B., Schwarzer, A., Kalinowski, J., Puhler, A., Eggeling, L. & Sahm, H. A functionally split pathway for lysine synthesis in *Corynebacterium glutamicium*. *Journal of bacteriology* **173**, 4510-6 (1991).

33. Pavelka, M. S., Jr., Weisbrod, T. R. & Jacobs, W. R., Jr. Cloning of the *dap*B gene, encoding dihydrodipicolinate reductase, from *Mycobacterium tuberculosis*. *Journal of bacteriology* **179**, 2777-82 (1997).

34. Work, E. A new naturally occurring amino acid. *Nature* **165**, 74 (1950).

35. Asselineau, J., Choucroun, N. & Lederer, E. On the chemical constitution of an antigenic lipo-polysaccharide extract of *Mycobacterium tuberculosis var. hominis*. *Biochimica et biophysica acta* **5**, 197-203 (1950).

36. Work, E. The isolation of α-ε-diaminopimelic acid from *Corynebacterium diphtheriae* and *Mycobacterium tuberculosis*. *The Biochemical journal* **49**, 17-23 (1951).

37. Meadow, P. & Work, E. Biosynthesis of diaminopimelic acid and lysine in *Escherichia coli*. 2. Incorporation of [14C] diaminopimelic acid, lysine and glucose. *The Biochemical journal* **72**, 400-7 (1959).

38. Hutton, C. A., Southwood, T. J. & Turner, J. J. Inhibitors of lysine biosynthesis as antibacterial agents. *Mini reviews in medicinal chemistry* **3**, 115-27 (2003).

39. Cox, R. J. The DAP pathway to lysine as a target for antimicrobial agents. *Natural product reports* **13**, 29-43 (1996).

40. Wietzerbin, J., Das, B. C., Petit, J. F., Lederer, E., Leyh-Bouille, M. & Ghuysen, J. M. Occurrence of D-alanyl-(D)-meso-diaminopimelic acid and meso-diaminopimelyl-meso-diaminopimelic acid interpeptide linkages in the peptidoglycan of Mycobacteria. *Biochemistry* **13**, 3471-6 (1974).

41. Cox, R. J., Sutherland, A. & Vederas, J. C. Bacterial diaminopimelate metabolism as a target for antibiotic design. *Bioorganic & medicinal chemistry letters* **8**, 843-71 (2000).

42. Girodeau, J. M., Agouridas, C., Masson, M., Pineau, R. & Le Goffic, F. The lysine pathway as a target for a new genera of synthetic antibacterial antibiotics? *Journal of medicinal chemistry* **29**, 1023-30 (1986).

43. Petit, J. F., Wietzerbin, J., Das, B. C. & Lederer, E. Chemical structure of the cell wall of *Mycobacterium tuberculosis var. bovis*, strain BCG. *Zeitschrift für Immunitätsforschung, experimentelle und klinische Immunologie* **149**, 118-25 (1975).

44. Consaul, S. A., Jacobs, W. R., Jr. & Pavelka, M. S., Jr. Extragenic suppression of the requirement for diaminopimelate in diaminopimelate auxotrophs of *Mycobacterium smegmatis*. *FEMS microbiology letters* **225**, 131-5 (2003).

45. Consaul, S. A., Wright, L. F., Mahapatra, S., Crick, D. C. & Pavelka, M. S., Jr. An unusual mutation results in the replacement of diaminopimelate with lanthionine in the peptidoglycan of a mutant strain of *Mycobacterium smegmatis*. *Journal of bacteriology* **187**, 1612-20 (2005).

46. Flores, A. R., Parsons, L. M. & Pavelka, M. S., Jr. Characterization of novel *Mycobacterium tuberculosis* and *Mycobacterium smegmatis* mutants hypersusceptible to β-lactam antibiotics. *Journal of bacteriology* **187**, 1892-900 (2005).

47. Acar, J. F. Antibiotic synergy and antagonism. *The Medical clinics of North America* **84**, 1391-406 (2000).

48. Monaghan, R. L. & Barrett, J. F. Antibacterial drug discovery--then, now and the genomics future. *Biochemical pharmacology* **71**, 901-9 (2006).

49. Zabriskie, T. M. & Jackson, M. D. Lysine biosynthesis and metabolism in fungi. *Natural product reports* **17**, 85-97 (2000).

50. Pavelka, M. S., Jr. & Jacobs, W. R., Jr. Comparison of the construction of unmarked deletion mutations in *Mycobacterium smegmatis*, *Mycobacterium bovis* bacillus Calmette-Guerin, and *Mycobacterium tuberculosis* H37Rv by allelic exchange. *Journal of bacteriology* **181**, 4780-9 (1999).

51. Blickling, S. & Knablein, J. Feedback inhibition of dihydrodipicolinate synthase enzymes by L-lysine. *Biological chemistry* **378**, 207-10 (1997).

52. Stahly, D. P. Dihydrodipicolinic acid synthase of *Bacillus licheniformis*. *Biochimica et biophysica acta* **191**, 439-51 (1969).

53. Acord, J. & Masters, M. Expression from the *Escherichia coli dap*A promoter is regulated by intracellular levels of diaminopimelic acid. *FEMS microbiology letters* **235**, 131-7 (2004).

54. Butour, J. L., Felenbok, B. & Patte, J. C. Synthesis of dihydrodipicolinate synthetase in *Escherichia coli* K12. *Annales de microbiologie* **125**, 459-62 (1974).

55. Richaud, F., Richaud, C., Ratet, P. & Patte, J. C. Chromosomal location and nucleotide sequence of the *Escherichia coli dap*A gene. *Journal of bacteriology* **166**, 297-300 (1986).

56. Marcel, T., Archer, J. A., Mengin-Lecreulx, D. & Sinskey, A. J. Nucleotide sequence and organization of the upstream region of the *Corynebacterium glutamicum lys*A gene. *Molecular microbiology* **4**, 1819-30 (1990).

57. Blickling, S., Renner, C., Laber, B., Pohlenz, H. D., Holak, T. A. & Huber, R. Reaction mechanism of *Escherichia coli* dihydrodipicolinate synthase investigated by X-ray crystallography and NMR spectroscopy. *Biochemistry* **36**, 24-33 (1997).

58. Dobson, R. C., Griffin, M. D., Jameson, G. B. & Gerrard, J. A. The crystal structures of native and (*S*)-lysine-bound dihydrodipicolinate synthase from *Escherichia coli* with improved resolution show new features of biological significance. *Acta crystallographica. Section D, Biological crystallography* **61**, 1116-24 (2005).

59. Galili, G. New insights into the regulation and functional significance of lysine metabolism in plants. *Annual review of plant biology* **53**, 27-43 (2002).

60. Galili, G. Regulation of lysine and threonine synthesis. *The Plant cell* **7**, 899-906 (1995).

61. Blickling, S., Beisel, H. G., Bozic, D., Knablein, J., Laber, B. & Huber, R. Structure of dihydrodipicolinate synthase of *Nicotiana sylvestris* reveals novel quaternary structure. *Journal of molecular biology* **274**, 608-21 (1997).

62. Mirwaldt, C., Korndorfer, I. & Huber, R. The crystal structure of dihydrodipicolinate synthase from *Escherichia coli* at 2.5 Å resolution. *Journal of molecular biology* **246**, 227-39 (1995).

63. Rao, A. S. Regulation of lysine and dipicolinic acid biosynthesis in *Bacillus brevis* ATCC 10068: significance of derepression of the enzymes during the change from vegetative growth to sporulation. *Archives of microbiology* **141**, 143-50 (1985).

64. Hoganson, D. A. & Stahly, D. P. Regulation of dihydrodipicolinate synthase during growth and sporulation of *Bacillus cereus*. *Journal of bacteriology* **124**, 1344-50 (1975).

65. Chatterjee, M., Chatterjee, S. P. & Banerjee, A. K. Dihydrodipicolinate synthase of lysine excreting and nonexcreting strains of *Bacillus-megaterium*. *Acta biotechnologica* **10**, 382-4 (1990).

66. Webster, F. H. & Lechowich, R. V. Partial purification and characterization of dihydrodipicolinic acid synthetase from sporulating *Bacillus megaterium*. *Journal of bacteriology* **101**, 118-26 (1970).

67. Selli, A., Crociani, F., Digioia, D., Fava, F., Crisetig, G. & Matteuzzi, D. Regulation of dihydrodipicolinate synthase and diaminopimelate decarboxylase activity in *Bacillus-stearothermophilus*. *The Italian journal of biochemistry* **43**, 29-35 (1994).

68. Yamakura, F., Ikeda, Y., Kimura, K. & Sasakawa, T. Partial purification and some properties of pyruvate-aspartic semialdehyde condensing enzyme from sporulating *Bacillus subtilis*. *Journal of biochemistry* **76**, 611-21 (1974).

69. Miyajima, R. & Shio, I. Regulation of aspartate family amino acid biosynthesis in *Brevibacterium flavum*. V. Properties of homoserine kinase. *Journal of biochemistry* **71**, 219-26 (1972).

70. Tosaka, O. & Takinami, K. Pathway and regulation of lysine biosynthesis in *Brevibacterium lactofermentum*. *Agricultural and biological chemistry* **42**, 95-100 (1978).

71. Rice, E. A., Bannon, G. A., Glenn, K. C., Jeong, S. S., Sturman, E. J. & Rydel, T. J. Characterization and crystal structure of lysine insensitive *Corynebacterium glutamicum* dihydrodipicolinate synthase (cDHDPS) protein. *Archives of biochemistry and biophysics* **480**, 111-21 (2008).

72. Cremer, J., Treptow, C., Eggeling, L. & Sahm, H. Regulation of enzymes of lysine biosynthesis in *Corynebacterium glutamicum*. *Journal of general microbiology* **134**, 3221-9 (1988).

73. Laber, B., Gomis-Ruth, F. X., Romao, M. J. & Huber, R. *Escherichia coli* dihydrodipicolinate synthase. Identification of the active site and crystallization. *The Biochemical journal* **288 (Pt 2)**, 691-5 (1992).

74. Cahyanto, M. N., Kawasaki, H., Nagashio, M., Fujiyama, K. & Seki, T. Regulation of aspartokinase, aspartate semialdehyde dehydrogenase, dihydrodipicolinate synthase and dihydrodipicolinate reductase in *Lactobacillus plantarum*. *Microbiology* **152**, 105-12 (2006).

75. Gunji, Y., Tsujimoto, N., Shimaoka, M., Ogawa-Miyata, Y., Sugimoto, S. & Yasueda, H. Characterization of the L-lysine biosynthetic pathway in the obligate methylotroph *Methylophilus methylotrophus*. *Bioscience, biotechnology, and biochemistry* **68**, 1449-60 (2004).

76. Bakhiet, N., Forney, F., Stahly, D. & Daniels, L. Lysine biosynthesis in *Methanobacterium thermoautotrophicum* is by the diaminopimelic acid pathway. *Current microbiology* **10**, 195-8 (1984).

77. Devenish, S. R., Huisman, F. H., Parker, E. J., Hadfield, A. T. & Gerrard, J. A. Cloning and characterisation of dihydrodipicolinate synthase from the pathogen *Neisseria meningitidis*. *Biochimica et biophysica acta* **1794**, 1168-74 (2009).

78. Hermann, M., Thevenet, N. J., Coudert-Maratier, M. M. & Vandecasteele, J. P. Consequences of lysine oversynthesis in *Pseudomonas* mutants insensitive to feedback inhibition. Lysine excretion or endogenous induction of a lysine-catabolic pathway. *European journal of biochemistry / FEBS* **30**, 100-6 (1972).

79. Tam, P. H., Phenix, C. P. & Palmer, D. R. MosA, a protein implicated in rhizopine biosynthesis in *Sinorhizobium meliloti* L5-30, is a dihydrodipicolinate synthase. *Journal of molecular biology* **335**, 393-7 (2004).

80. Girish, T. S., Sharma, E. & Gopal, B. Structural and functional characterization of *Staphylococcus aureus* dihydrodipicolinate synthase. *FEBS letters* **582**, 2923-30 (2008).

81. Burgess, B. R., Dobson, R. C., Bailey, M. F., Atkinson, S. C., Griffin, M. D., Jameson, G. B., Parker, M. W., Gerrard, J. A. & Perugini, M. A. Structure and evolution of a novel dimeric enzyme from a clinically-important bacterial pathogen. *The Journal of biological chemistry* **283**, 27598-603 (2008).

82. Gilboe, D. P., Friede, J. D. & Henderson, L. M. Effect of hydroxylysine on the biosynthesis of lysine in *Streptococcus faecalis*. *Journal of bacteriology* **95**, 856-63 (1968).

83. Mendelovitz, S. & Aharonowitz, Y. Regulation of cephamycin C synthesis, aspartokinase, dihydrodipicolinic acid synthetase, and homoserine dehydrogenase by aspartic acid family amino acids in *Streptomyces clavuligerus*. *Antimicrobial agents and chemotherapy* **21**, 74-84 (1982).

84. Wolterink-van Loo, S., Levisson, M., Cabrieres, M. C., Franssen, M. C. & van der Oost, J. Characterization of a thermostable dihydrodipicolinate synthase from *Thermoanaerobacter tengcongensis*. *Extremophiles* **12**, 461-9 (2008).

85. Pearce, F. G., Perugini, M. A., McKerchar, H. J. & Gerrard, J. A. Dihydrodipicolinate synthase from *Thermotoga maritima*. *The Biochemical journal* **400**, 359-66 (2006).

86. Bey, P. Mechanism-based enzyme inhibitors as an approach to drug design in *Enzymes as targets for drug design* (eds. Palfreyman, M. G., McCann, P. P., Lovenberg, W., Temple, J. G. & Sjoerdsma, A.) 59-83 (Academic Press, Sydney, 1989).

87. Dobson, R. C., Valegard, K. & Gerrard, J. A. The crystal structure of three site-directed mutants of *Escherichia coli* dihydrodipicolinate synthase: further evidence for a catalytic triad. *Journal of molecular biology* **338**, 329-39 (2004).

88. Dobson, R. C., Griffin, M. D., Roberts, S. J. & Gerrard, J. A. Dihydrodipicolinate synthase (DHDPS) from *Escherichia coli* displays partial mixed inhibition with respect to its first substrate, pyruvate. *Biochimie* **86**, 311-5 (2004).

89. Karsten, W. E. Dihydrodipicolinate synthase from *Escherichia coli*: pH dependent changes in the kinetic mechanism and kinetic mechanism of allosteric inhibition by L-lysine. *Biochemistry* **36**, 1730-9 (1997).

90. Cornish-Bowden, A. *Fundamentals of enzyme kinetics* (Portland Press, London, 1995).

91. Cornish-Bowden, A. *Fundamentals of enzyme kinetics* (Butterworths, London ; Boston, 1979).

92. Dereppe, C., Bold, G., Ghisalba, O., Ebert, E. & Schar, H. P. Purification and characterization of dihydrodipicolinate synthase from pea. *Plant physiology* **98**, 813-21 (1992).

93. Frisch, D. A., Gengenbach, B. G., Tommey, A. M., Sellner, J. M., Somers, D. A. & Myers, D. E. Isolation and characterization of dihydrodipicolinate synthase from maize. *Plant physiology* **96**, 444-52 (1991).

94. Kumpaisal, R., Hashimoto, T. & Yamada, Y. Purification and characterization of dihydrodipicolinate synthase from wheat suspension cultures. *Plant physiology* **85**, 145-51 (1987).

95. Borthwick, E. B., Connell, S. J., Tudor, D. W., Robins, D. J., Shneier, A., Abell, C. & Coggins, J. R. *Escherichia coli* dihydrodipicolinate synthase: characterization of the imine intermediate and the product of bromopyruvate treatment by electrospray mass spectrometry. *The Biochemical journal* **305 (Pt 2)**, 521-4 (1995).

96. Dobson, R. C., Griffin, M. D., Devenish, S. R., Pearce, F. G., Hutton, C. A., Gerrard, J. A., Jameson, G. B. & Perugini, M. A. Conserved main-chain peptide distortions: A proposed role for Ile203 in catalysis by dihydrodipicolinate synthase. *Protein science* **17**, 2080-90 (2008).

97. Phenix, C. P., Nienaber, K., Tam, P. H., Delbaere, L. T. & Palmer, D. R. Structural, functional and calorimetric investigation of MosA, a dihydrodipicolinate synthase from *Sinorhizobium meliloti* l5-30, does not support involvement in rhizopine biosynthesis. *Chembiochem: a European journal of chemical biology* **9**, 1591-602 (2008).

98. Coulter, C. V., Gerrard, J. A., Kraunsoe, J. A. E. & Pratt, A. J. *Escherichia coli* dihydrodipicolinate synthase and dihydrodipicolinate reductase: kinetic and inhibition studies of two putative herbicide targets. *Pesticide science* **55**, 887-95 (1999).

99. Wallsgrove, R. M. & Mazelis, M. Spinach leaf dihydrodipicolinate synthase - partial-purification and characterization. *Phytochemistry* **20**, 2651-5 (1981).

100. Mazelis, M., Whatley, F. R. & Whatley, J. The enzymology of lysine biosynthesis in higher plants. The occurrence, characterization and some regulatory properties of dihydrodipicolinate synthase. *FEBS letters* **84**, 236-40 (1977).

101. Halling, S. M. & Stahly, D. P. Dihydrodipicolinic acid synthase of *Bacillus licheniformis*. Quaternary structure, kinetics, and stability in the presence of sodium chloride and substrates. *Biochimica et biophysica acta* **452**, 580-96 (1976).

102. Tosaka, O., Ishihara, M., Morinaga, Y. & Takinami, K. Mode of conversion of asparto β-semialdehyde to L-threonine and L-lysine in *Brevibacterium lactofermentum*. *Agricultural and biological chemistry* **43**, 265-70 (1979).

103. Dobson, R. C., Gerrard, J. A. & Pearce, F. G. Dihydrodipicolinate synthase is not inhibited by its substrate, (*S*)-aspartate β-semialdehyde. *The Biochemical journal* **377**, 757-62 (2004).

104. Shedlarski, J. G. & Gilvarg, C. The pyruvate-aspartic semialdehyde condensing enzyme of *Escherichia coli*. *The Journal of biological chemistry* **245**, 1362-73 (1970).

105. Grazi, E., Meloche, H., Martinez, G., Wood, W. A. & Horecker, B. L. Evidence for Schiff base formation in enzymatic aldol condensations. *Biochemical and biophysical research communications* **10**, 4-10 (1963).

106. Babbitt, P. C. & Gerlt, J. A. Understanding enzyme superfamilies. Chemistry as the fundamental determinant in the evolution of new catalytic activities. *The Journal of biological chemistry* **272**, 30591-4 (1997).

107. Barbosa, J. A., Smith, B. J., DeGori, R., Ooi, H. C., Marcuccio, S. M., Campi, E. M., Jackson, W. R., Brossmer, R., Sommer, M. & Lawrence, M. C. Active site modulation in the N-acetylneuraminate lyase sub-family as revealed by the structure of the inhibitor-complexed *Haemophilus influenzae* enzyme. *Journal of molecular biology* **303**, 405-21 (2000).

108. Joerger, A. C., Mayer, S. & Fersht, A. R. Mimicking natural evolution *in vitro*: an N-acetylneuraminate lyase mutant with an increased dihydrodipicolinate synthase activity. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 5694-9 (2003).

109. Izard, T., Lawrence, M. C., Malby, R. L., Lilley, G. G. & Colman, P. M. The three-dimensional structure of N-acetylneuraminate lyase from *Escherichia coli*. *Structure (London, England: 1993)* **2**, 361-9 (1994).

110. Lawrence, M. C., Barbosa, J. A., Smith, B. J., Hall, N. E., Pilling, P. A., Ooi, H. C. & Marcuccio, S. M. Structure and mechanism of a sub-family of enzymes related to N-acetylneuraminate lyase. *Journal of molecular biology* **266**, 381-99 (1997).

111. Theodossis, A., Walden, H., Westwick, E. J., Connaris, H., Lamble, H. J., Hough, D. W., Danson, M. J. & Taylor, G. L. The structural basis for substrate promiscuity in 2-keto-3-deoxygluconate aldolase from the Entner-Doudoroff pathway in *Sulfolobus solfataricus*. *The Journal of biological chemistry* **279**, 43886-92 (2004).

112. Manicka, S., Peleg, Y., Unger, T., Albeck, S., Dym, O., Greenblatt, H. M., Bourenkov, G., Lamzin, V., Krishnaswamy, S. & Sussman, J. L. Crystal structure of YagE, a putative DHDPS-like protein from *Escherichia coli* K12. *Proteins* **71**, 2102-8 (2008).

113. Devenish, S. R., Gerrard, J. A., Jameson, G. B. & Dobson, R. C. The high-resolution structure of dihydrodipicolinate synthase from *Escherichia coli* bound to its first substrate, pyruvate. *Acta crystallographica. Section F, Structural biology and crystallization communications* **64**, 1092-5 (2008).

114. Tudor, D. W., Lewis, T. & Robins, D. J. Synthesis of the trifluoroacetate salt of aspartic-acid β-semialdehyde, an Intermediate in the biosynthesis of L-lysine, L-threonine, and L-methionine. *Synthesis*1061-2 (1993).

115. Dobson, R. C., Devenish, S. R., Turner, L. A., Clifford, V. R., Pearce, F. G., Jameson, G. B. & Gerrard, J. A. Role of arginine 138 in the catalysis and regulation of *Escherichia coli* dihydrodipicolinate synthase. *Biochemistry* **44**, 13007-13 (2005).

116. Turner, J. J., Gerrard, J. A. & Hutton, C. A. Heterocyclic inhibitors of dihydrodipicolinate synthase are not competitive. *Bioorganic & medicinal chemistry* **13**, 2133-40 (2005).

117. Boughton, B. A., Dobson, R. C., Gerrard, J. A. & Hutton, C. A. Conformationally constrained diketopimelic acid analogues as inhibitors of dihydrodipicolinate synthase. *Bioorganic & medicinal chemistry letters* **18**, 460-3 (2008).

118. Boughton, B. A., Griffin, M. D., O'Donnell, P. A., Dobson, R. C., Perugini, M. A., Gerrard, J. A. & Hutton, C. A. Irreversible inhibition of dihydrodipicolinate synthase by 4-oxo-heptenedioic acid analogues. *Bioorganic & medicinal chemistry* **16**, 9975-83 (2008).

119. Kumpaisal, R., Hashimoto, T. & Yamada, Y. Inactivation of wheat dihydrodipicolinate synthase by 3-bromopyruvate. *Agricultural and biological chemistry* **53**, 355-9 (1989).

120. Coulter, C. V., Gerrard, J. A., Kraunsoe, J. A. E., Moore, D. J. & Pratt, A. J. *(S)*-Aspartate semi-aldehyde: Synthetic and structural studies. *Tetrahedron* **52**, 7127-36 (1996).

121. Couper, L., McKendrick, J. E., Robins, David J. & Chrystal, E. J. T. Pyridine and piperidine derivatives as inhibitors of dihydrodipicolinic acid synthase, a key enzyme in the diaminopimelate pathway to L-lysine. *Bioorganic & medicinal chemistry letters* **4**, 2267-72 (1994).

122. Mitsakos, V., Dobson, R. C., Pearce, F. G., Devenish, S. R., Evans, G. L., Burgess, B. R., Perugini, M. A., Gerrard, J. A. & Hutton, C. A. Inhibiting dihydrodipicolinate synthase across species: towards specificity for pathogens? *Bioorganic & medicinal chemistry letters* **18**, 842-4 (2008).

123. Gerrard, J. A., Hutton, C. A. & Perugini, M. A. Inhibiting protein-protein interactions as an emerging paradigm for drug discovery. *Mini reviews in medicinal chemistry* **7**, 151-7 (2007).

124. Traut, T. W. Dissociation of enzyme oligomers: a mechanism for allosteric regulation. *Critical reviews in biochemistry and molecular biology* **29**, 125-63 (1994).

125. Beernink, P. T. & Tolan, D. R. Disruption of the aldolase A tetramer into catalytically active monomers. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 5374-9 (1996).

126. Klotz, I. M., Langerman, N. R. & Darnall, D. W. Quaternary structure of proteins. *Annual review of biochemistry* **39**, 25-62 (1970).

127. Arkin, M. R. & Wells, J. A. Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. *Natural reviews. Drug discovery* **3**, 301-17 (2004).

128. Gadek, T. R. & Nicholas, J. B. Small molecule antagonists of proteins. *Biochemical pharmacology* **65**, 1-8 (2003).

129. Fry, D. C. Protein-protein interactions as targets for small molecule drug discovery. *Biopolymers* **84**, 535-52 (2006).

130. Nooren, I. M. & Thornton, J. M. Diversity of protein-protein interactions. *The EMBO journal* **22**, 3486-92 (2003).

131. Archakov, A. I., Govorun, V. M., Dubanov, A. V., Ivanov, Y. D., Veselovsky, A. V., Lewi, P. & Janssen, P. Protein-protein interactions as a target for drugs in proteomics. *Proteomics* **3**, 380-91 (2003).

132. Wells, J. A. & McClendon, C. L. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* **450**, 1001-9 (2007).

133. Park, C. M., Bruncko, M., Adickes, J., Bauch, J., Ding, H., Kunzer, A., Marsh, K. C., Nimmer, P., Shoemaker, A. R., Song, X., Tahir, S. K., Tse, C., Wang, X., Wendt, M. D., Yang, X., Zhang, H., Fesik, S. W., Rosenberg, S. H. & Elmore, S. W. Discovery of an orally bioavailable small molecule inhibitor of prosurvival B-cell lymphoma 2 proteins. *Journal of medicinal chemistry* **51**, 6902-15 (2008).

134. Bogan, A. A. & Thorn, K. S. Anatomy of hot spots in protein interfaces. *Journal of molecular biology* **280**, 1-9 (1998).

135. Clackson, T. & Wells, J. A. A hot spot of binding energy in a hormone-receptor interface. *Science* **267**, 383-6 (1995).

136. Schwarz, M., Meade, G., Stoll, P., Ylanne, J., Bassler, N., Chen, Y. C., Hagemeyer, C. E., Ahrens, I., Moran, N., Kenny, D., Fitzgerald, D., Bode, C. & Peter, K. Conformation-specific blockade of the integrin GPIIb/IIIa: a novel antiplatelet strategy that selectively targets activated platelets. *Circulation research* **99**, 25-33 (2006).

137. McMillan, K., Adler, M., Auld, D. S., Baldwin, J. J., Blasko, E., Browne, L. J., Chelsky, D., Davey, D., Dolle, R. E., Eagen, K. A., Erickson, S., Feldman, R. I., Glaser, C. B., Mallari, C., Morrissey, M. M., Ohlmeyer, M. H., Pan, G., Parkinson, J. F., Phillips, G. B., Polokoff, M. A., Sigal, N. H., Vergona, R., Whitlow, M., Young, T. A. & Devlin, J. J. Allosteric inhibitors of inducible nitric oxide synthase dimerization discovered via combinatorial chemistry. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 1506-11 (2000).

138. He, M. M., Smith, A. S., Oslob, J. D., Flanagan, W. M., Braisted, A. C., Whitty, A., Cancilla, M. T., Wang, J., Lugovskoy, A. A., Yoburn, J. C., Fung, A. D., Farrington, G., Eldredge, J. K., Day, E. S., Cruz, L. A., Cachero, T. G., Miller, S. K., Friedman, J. E., Choong, I. C. & Cunningham, B. C. Small-molecule inhibition of TNF-α. *Science* **310**, 1022-5 (2005).

139. Fletcher, S. & Hamilton, A. D. Targeting protein-protein interactions by rational design: mimicry of protein surfaces. *Journal of the Royal Society, Interface/the Royal Society* **3**, 215-33 (2006).

140. Sillerud, L. O. & Larson, R. S. Design and structure of peptide and peptidomimetic antagonists of protein-protein interaction. *Current protein & peptide science* **6**, 151-69 (2005).

141. Zhang, Z. Y., Poorman, R. A., Maggiora, L. L., Heinrikson, R. L. & Kezdy, F. J. Dissociative inhibition of dimeric enzymes. Kinetic characterization of the inhibition of HIV-1 protease by its COOH-terminal tetrapeptide. *The Journal of biological chemistry* **266**, 15591-4 (1991).

142. Schramm, H. J., Nakashima, H., Schramm, W., Wakayama, H. & Yamamoto, N. HIV-1 reproduction is inhibited by peptides derived from the N- and C-termini of HIV-1 protease. *Biochemical and biophysical research communications* **179**, 847-51 (1991).

143. Boggetto, N. & Reboud-Ravaux, M. Dimerization inhibitors of HIV-1 protease. *Biological chemistry* **383**, 1321-4 (2002).

144. Bannwarth, L., Rose, T., Dufau, L., Vanderesse, R., Dumond, J., Jamart-Gregoire, B., Pannecouque, C., De Clercq, E. & Reboud-Ravaux, M. Dimer disruption and monomer sequestration by alkyl tripeptides are successful strategies for inhibiting wild-type and multidrug-resistant mutated HIV-1 proteases. *Biochemistry* **48**, 379-87 (2009).

145. Perugini, M. A., Griffin, M. D., Smith, B. J., Webb, L. E., Davis, A. J., Handman, E. & Gerrard, J. A. Insight into the self-association of key enzymes from pathogenic species. *European biophysics journal* **34**, 469-76 (2005).

146. Blagova, E., Levdikov, V., Milioti, N., Fogg, M. J., Kalliomaa, A. K., Brannigan, J. A., Wilson, K. S. & Wilkinson, A. J. Crystal structure of dihydrodipicolinate synthase (BA3935) from *Bacillus anthracis* at 1.94 A resolution. *Proteins* **62**, 297-301 (2006).

147. Dobson, R. C. J. Investigating the catalytic and regulatory mechanisms of dihydrodipicolinate synthase. *D. Phil Thesis*, Univeristy of Canterbury. (2003).

148. Pearce, F. G., Dobson, R. C., Weber, A., Lane, L. A., McCammon, M. G., Squire, M. A., Perugini, M. A., Jameson, G. B., Robinson, C. V. & Gerrard, J. A. Mutating the tight-dimer interface of dihydrodipicolinate synthase disrupts the enzyme quaternary structure: toward a monomeric enzyme. *Biochemistry* **47**, 12108-17 (2008).

149. Kefala, G., Evans, G. L., Griffin, M. D., Devenish, S. R., Pearce, F. G., Perugini, M. A., Gerrard, J. A., Weiss, M. S. & Dobson, R. C. Crystal structure and kinetic study of dihydrodipicolinate synthase from *Mycobacterium tuberculosis*. *The Journal of biological chemistry* **411**, 351-60 (2008).

150. Leduc, Y. A., Phenix, C. P., Puttick, J., Nienaber, K., Palmer, D. R. & Delbaere, L. T. Crystallization, preliminary X-ray diffraction and structure solution of MosA, a dihydrodipicolinate synthase from *Sinorhizobium meliloti* L5-30. *Acta crystallographica. Section F, Structural biology and crystallization communications* **62**, 49-51 (2006).

151. Pearce, F. G., Perugini, M. A., McKerchar, H. J. & Gerrard, J. A. Dihydrodipicolinate synthase from Thermotoga maritima. *Biochem J* (2006).

152. Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. The Protein Data Bank: a computer-based archival file for macromolecular structures. *Journal of molecular biology* **112**, 535-42 (1977).

153. DeLano, W. L. The PyMOL molecular graphics system. (DeLano Scientific, San Carlos, CA, USA, 2006).

154. Nishida, H., Nishiyama, M., Kobashi, N., Kosuge, T., Hoshino, T. & Yamane, H. A prokaryotic gene cluster involved in synthesis of lysine through the amino adipate pathway: a key to the evolution of amino acid biosynthesis. *Genome research* **9**, 1175-83 (1999).

155. Kobashi, N., Nishiyama, M. & Tanokura, M. Aspartate kinase-independent lysine synthesis in an extremely thermophilic bacterium, *Thermus thermophilus*: lysine is synthesized via alpha-aminoadipic acid not via diaminopimelic acid. *Journal of bacteriology* **181**, 1713-8 (1999).

156. Kosuge, T. & Hoshino, T. Lysine is synthesized through the α-aminoadipate pathway in *Thermus thermophilus*. *FEMS microbiology letters* **169**, 361-7 (1998).

157. Ghislain, M., Frankard, V. & Jacobs, M. A dinucleotide mutation in dihydrodipicolinate synthase of *Nicotiana sylvestris* leads to lysine overproduction. *The Plant journal: for cell and molecular biology* **8**, 733-43 (1995).

158. Griffin, M. D. W. Why is dihydrodipicolinate synthase a tetramer? *D. Phil Thesis*, Univeristy of Canterbury. (2005).

159. Wang, M., Yafremava, L. S., Caetano-Anolles, D., Mittenthal, J. E. & Caetano-Anolles, G. Reductive evolution of architectural repertoires in proteomes and the birth of the tripartite world. *Genome research* **17**, 1572-85 (2007).

160. Purich, D. L. *Contemporary enzyme kinetics and mechanism* (Academic Press, San Diego, 1996).

161. Banner, D. W., Bloomer, A. C., Petsko, G. A., Phillips, D. C., Pogson, C. I., Wilson, I. A., Corran, P. H., Furth, A. J., Milman, J. D., Offord, R. E., Priddle, J. D. & Waley, S. G. Structure of chicken muscle triose phosphate isomerase determined crystallographically at 2.5 angstrom resolution using amino acid sequence data. *Nature* **255**, 609-14 (1975).

162. Devenish, S. R. & Gerrard, J. A. The role of quaternary structure in $(\beta/\alpha)_8$-barrel proteins: evolutionary happenstance or a higher level of structure-function relationships? *Organic and biomolecular chemistry* **7**, 833-9 (2009).

Chapter Two

Purification & characterization of wild-type DHDPS from *M. tuberculosis*

2.1 Introduction

The aim of this thesis is to examine DHDPS from *M. tuberculosis* as a putative drug target. The first step was to fully characterize the wild-type enzyme, which required the production and purification of milligram quantities of active *M. tuberculosis* DHDPS. Our collaborators, at EMBL Hamburg Outstation, Dr Manfred Weiss and Dr Georgia Kefala, cloned the *dap*A gene for *M. tuberculosis* DHDPS, from the bacterial strain H37Rv, into a plasmid vector, for expression in *E. coli*.[1] This vector, pETM-11, coded for a cleavable N-terminal His_6-tag allowing for purification using affinity chromatography.[2]

Two methods, a colorimetric and a coupled assay, were used to monitor activity throughout the purification procedure to ensure functionally active *M. tuberculosis* DHDPS was obtained. The coupled assay required the additional purification of a coupling enzyme, DHDPR, and was used for kinetic characterization. Steady-state kinetics and biophysical methods were used to study the catalytic behaviour and solution properties of *M. tuberculosis* DHDPS. The results were considered in the context of the solid-state structure determined by X-ray crystallography by our collaborators, Weiss and Kefala, and this characterization of *M. tuberculosis* DHDPS has been published in *Biochemical Journal* as a joint co-first author manuscript with Kefala **(**see **Appendix J)**: "Crystal structure and kinetic study of dihydrodipicolinate synthase from *Mycobacterium tuberculosis*."[3]

As discussed in chapter one, inhibiting the activity of DHDPS by disrupting quaternary structure is the approach to drug design primarily considered in this thesis. However, the effect of (*S*)-lysine on the activity of *M. tuberculosis* DHDPS was also of interest. *E. coli* DHDPS is weakly inhibited by (*S*)-lysine and contains binding sites for (*S*)-lysine (discussed in chapter one, sections 1.2.4 and 1.5.4, respectively). Several inhibitors designed to interact at the active site of *E. coli* DHDPS were found not potent enough for biological use and their mode of inhibition suggested they were interacting at the (*S*)-lysine binding site instead of the

active site.[4] Associated with this research, inhibitory activity of compounds designed to interact at the active site were examined for DHDPS from several pathogenic species including *M. tuberculosis*, and showed species selective inhibition, perhaps reflecting the absence or presence of (*S*)-lysine binding sites. This work has been published in *Bioorganic and Biomedical Chemistry Letters* in a brief communication **(**see **Appendix K)** on which I am a co-author, since I supplied the *M. tuberculosis* DHDPS for this collaborative study.[5]

2.2 Methods for monitoring DHDPS activity

The discoverers of DHDPS, Yugari and Gilvarg, developed three different assay systems to detect its activity.[6] The coupled assay was described as the method of choice for evaluating the properties of purified DHDPS, whereas the simpler assay systems, which use imidazole or *o*-aminobenzaldehyde, were deemed primarily useful for detecting enzyme activity during purification.[6,7]

2.2.1 The *o*-aminobenzaldehyde assay

o-Aminobenzaldehyde has been shown to interact with cyclized nitrogen containing compounds, such as the product of the reaction catalyzed by DHDPS, to form salts with a characteristic yellow/orange colour.[6] A reddish purple chromophore is produced after a 20 to 30 minute incubation of *o*-aminobenzaldehyde with DHDPS and its substrates.[6,7] As the exact chemical nature of the coloured adduct and the rate determining step in its formation are unclear, the assay was suggested to be mainly useful for determining DHDPS activity in crude extracts.[6] Nevertheless, the *o*-aminobenzaldehyde assay, monitoring the increase in absorbance at 540 nm after a 30 minute lag time, has been used by a number of groups to measure initial-rates for kinetic studies,[8-17] making the key assumption that the enzyme-catalyzed reaction, rather than chromophore formation, is the rate limiting step. While this assumption may or may not be valid, the assay does have several favourable features. Its high specificity, high sensitivity, lack of interfering side reactions, relatively few components and colorimetric nature has made it amenable to scaling down for semi-quantitative monitoring of DHDPS activity throughout the purification procedure,[7] and, thus, the *o*-aminobenzaldehyde assay was used for this purpose throughout this work.

2.2.2 The imidazole assay

The original method recommended for monitoring DHDPS activity during purification was the incubation of the enzyme and its substrates in imidazole buffer at pH 7.4, which gives rise to a peak at 270 nm after a 10 minute lag time.[7] Again, the exact nature of the chromophore has not been determined; however, it must be an isomeric or degraded form of physiologically active product, since the increase in absorbance at 270 nm corresponds to a decrease in the activity of the next enzyme in the (*S*)-lysine/DAP biosynthetic pathway, DHDPR.[6] Consequently, the imidazole buffer assay, like the *o*-aminobenzaldehyde assay, was not recommended for kinetic analysis,[6] but due to its simplicity and ease of use it has been used for such purpose by several workers.[18-21]

2.2.3 The coupled assay

The development of the coupled assay was based on the simple principle that the rate of an enzymatic reaction causing no appreciable spectrophotometric change can be monitored through the consumption of its product by a second observable reaction.[22-24] When the product from the reaction of interest reacts rapidly and does not accumulate then the measured rate of the second reaction corresponds to the rate of the first reaction.[22]

Conveniently, DHDPR, the next enzyme in the DAP/(*S*)-lysine biosynthetic pathway, both consumes the product of DHDPS and utilizes the cofactor NADPH **(Figure 2.1)**.[25] In its reduced state, NADPH has a strong absorbance peak at 340 nm, which disappears after oxidation to $NADP^+$.[22] Many spectrophotometric assays take advantage of this naturally occurring chromophore with its large molar extinction coefficient (ε_{340nm} = 6220 $M^{-1}.cm^{-1}$).[26] With an excess of DHDPR, the decrease in absorbance at 340 nm corresponds to DHDPS activity,[6] which is easily quantifiable since the product of the reaction of DHDPS has a 1:1 stoichiometric relationship to the oxidation of NADPH **(Figure 2.1)**.

The chief drawback of the coupled assay is that it requires the additional purification of DHDPR, so, unsurprisingly, before advances in protein purification such as recombinant DNA technology, many early investigators used the other assay methods for kinetic analysis (see chapter one, section 1.3.1, table 1.2). However, more recent work, including several comprehensive kinetic analysis of DHDPS from *E. coli*, have utilized the coupled assay.[27-29]

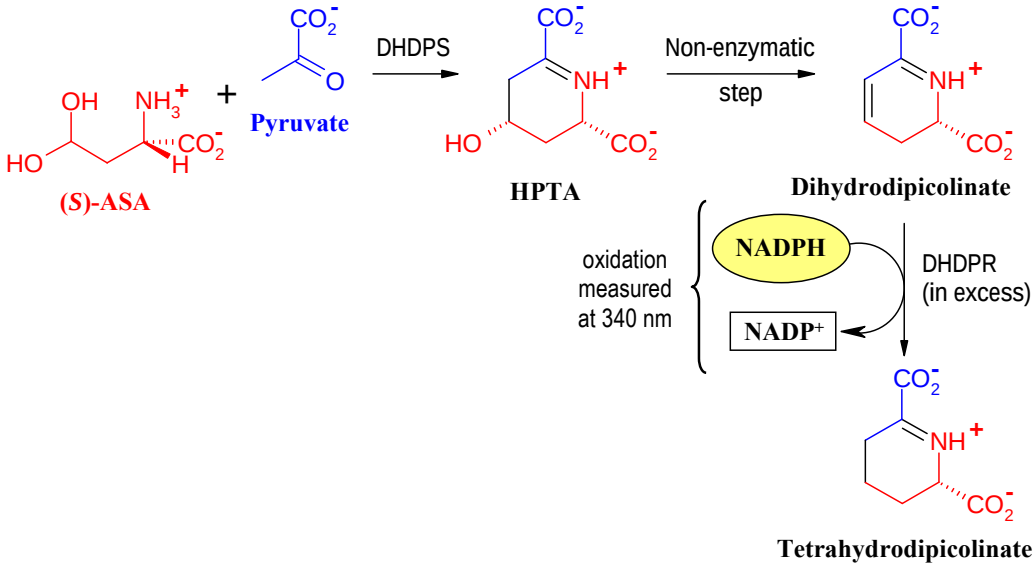


Figure 2.1: The preferred assay for kinetic analysis couples the observable reaction of DHDPR oxidizing NADPH to the activity of DHDPS.[4]

In this work, the coupled assay was used for kinetic study, based on an optimized procedure for *E. coli* DHDPS.[30] The coupled assay was also used for quantitative monitoring of enzyme activity and purity throughout the purification procedure; however, these results were less accurate due to the possible presence of NADPH-utilizing enzymes during various purification steps, and therefore it was used in tandem with the *o*-aminobenzaldehyde assay.

2.3 Enzyme over-expression & purification

In-depth understanding of gene expression in *E. coli* has allowed the development of vectors and bacterial strains to over-express almost any gene product.[31] Unfortunately, the resultant over-expression of proteins in *E. coli* does not necessarily lead to simple protein purification. Many recombinant proteins misfold and are insoluble when expressed in *E. coli*, as opposed to their native microorganism.[32] In fact, the TB Structural Genomics Consortium (http://www.doe-mbi.ucla.edu/TB) has found approximately half of their targeted *M. tuberculosis* proteins were insoluble when expressed in *E. coli*.[32] A more generalized look at results from several large structural genomic centres, shows that more than half of the recombinant proteins overproduced in *E. coli* form insoluble aggregates as opposed to correctly folded proteins.[33,34] The yield of *M. tuberculosis* DHDPS purified as a recombinant protein in *E. coli* was a limiting factor throughout this research, due to the majority being

expressed in insoluble form. The *dap*A gene for *M. tuberculosis* DHDPS was cloned into a vector, called pETM-11 **(Appendix A)** by our collaborator, Kefala,[1] resulting in the plasmid construct pMTB02. The plasmid construct, pMTB02, codes for *M. tuberculosis* DHDPS as a fused protein with a linker sequence and His_6-tag **(**details in **Appendix C)**.[1]

2.3.1 Original purification protocol developed for *M. tuberculosis* DHDPS

Initially, over-expression and purification of *M. tuberculosis* DHDPS was based on the method developed by Kefala.[1] LB broth medium was inoculated with cells transformed by the plasmid, pMTB02 **(Appendix A)**. These cultures were supplemented with kanamycin and chloramphenicol, and over-expression was induced after several hours by isopropyl thiogalactoside (IPTG). Following induction, the cells were grown overnight at room temperature and then harvested by centrifugation and re-suspended in extraction buffer. The extraction buffer contained tris(hydroxymethyl)methylamine (Tris.HCl), NaCl, imidazole, glycerol, β-mercaptoethanol (β-ME), BugBuster™ (Novagen) and EDTA-free protease inhibitor (Roche). Crude cell-free lysates were produced by extensive ultrasonication, for 20 minutes in 25 s pulses at 4 °C, followed by clarification using centrifugation. Over-expression of His_6-tagged *M. tuberculosis* DHDPS was apparent in the crude cell-free lysate, as shown by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) with Coomassie blue staining **(Figure 2.2, lane 2)**, but the majority of protein was lost during centrifugation, 60 minutes at 10000 rpm and 4 °C **(Figure 2.2, lane 3)**.

Purification using His_6-tag affinity chromatography involved several steps and included washing with three different buffers, which was found to improve protein purity (details in chapter six, section 6.4.3). The peak fractions were pooled and analyzed by SDS-PAGE **(Figure 2.2, lane 4)**, which showed His_6-tagged *M. tuberculosis* DHDPS with several additional protein bands reflecting low level impurities. This purification method gave a yield of 1.6 mg of His_6-tagged protein per L. Initially, the His_6-tagged *M. tuberculosis* DHDPS was kinetically characterized **(Appendix B)**, as cleavage of the His_6-tag required additional steps and therefore led to loss of protein yield. However, as cleavage of the His_6-tag was observed to increase enzyme activity, subsequent work, therefore, analyzed *M. tuberculosis* DHDPS which had been treated to remove the affinity tag.

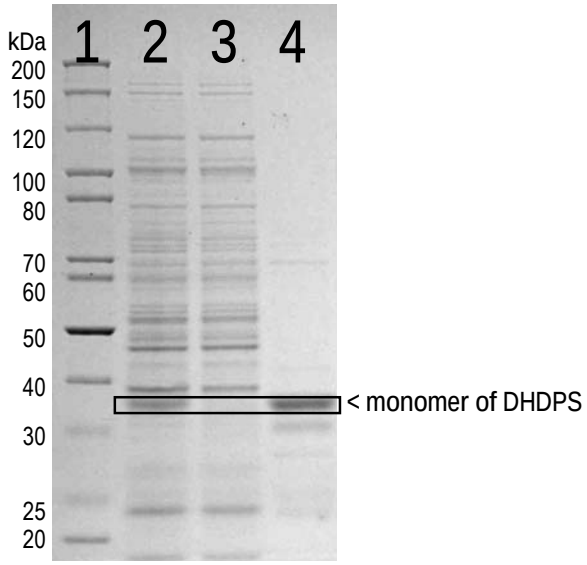


Figure 2.2: **The purification of recombinant *M. tuberculosis* DHDPS based on the originally developed protocol,[1] examined using denaturing SDS-PAGE, which separates based on subunit molecular weight (MW). *M. tuberculosis* DHDPS was identified as the 34 kDa band (black rectangle) using the MW marker, in lane 1. Protein preparations from purification stages were loaded as follows: 2 - crude extract, 3 - supernatant after centrifugation, 4 - eluted from His_6-tag affinity column.**

In order to remove the His_6-tag, fractions were incubated overnight with TEV protease **(**purchased or purified,[35] details in **Appendix C)** in the presence of 5 mM ethylenediamine tetraacetic acid (EDTA) and 1 mM dithiothreitol (DTT), followed by passing the cleaved protein through the His_6-tag affinity column. Protein purity was increased by gel-filtration chromatography, which further decreased protein yield, and in this work resulted in 0.25 mg of purified non-tagged protein per L of culture. This is lower yield than that reported by Kefala of 1 mg of *M. tuberculosis* DHDPS from 1 L of culture,[1] which accounts for only 1 % of the total protein expressed. Consequently, many variables were explored in an effort to optimize yield, including extraction techniques, host strains, growth temperature and length of induced expression. The most significant results are discussed in the following section.

2.3.2 Optimization of over-expression & purification of *M. tuberculosis* DHDPS

Centrifugation

The first variable explored for optimization was centrifugation. The effect of centrifugation for 5 minutes on crude lysate at low speeds of 1000, 2000, 3000, and 4000 rpm, was analyzed

by SDS-PAGE and the *o*-aminobenzaldehyde assay **(Figure 2.3)**. As the speed increased, the supernatant showed a marked decrease in the amount of His_6-tagged *M. tuberculosis* DHDPS (**Figure 2.3A, lanes 2-5)** and DHDPS activity **(Figure 2.3B),** and there was a corresponding increase in the amount of *M. tuberculosis* DHDPS in the pellet **(Figure 2.3A, lanes 7-10)**.

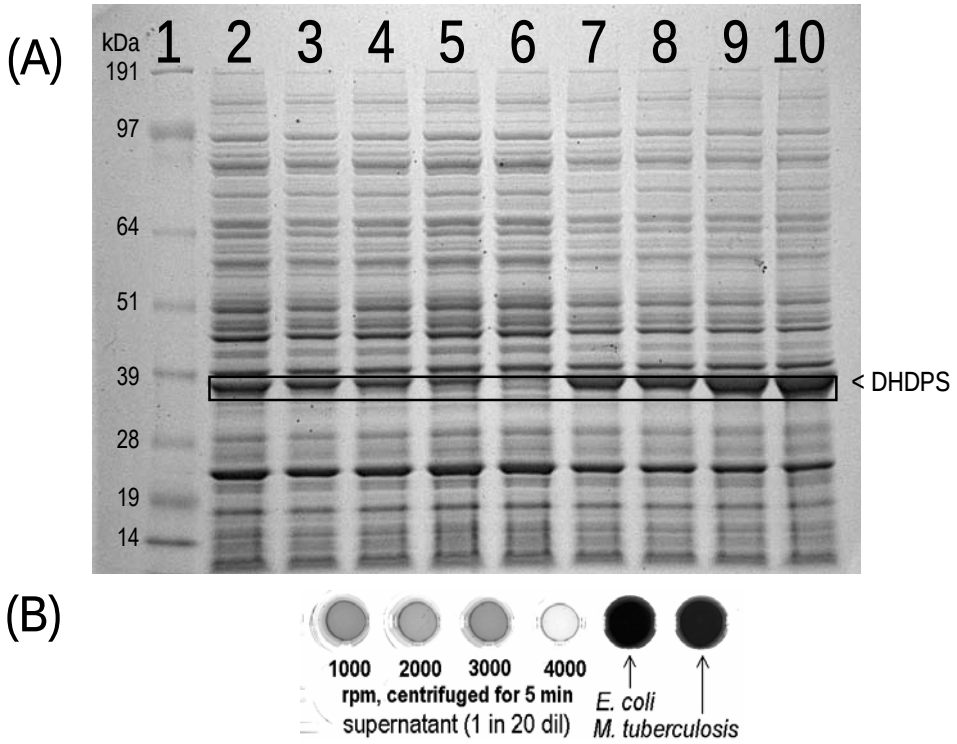


Figure 2.3: **The effect of varied low speed centrifugation for 5 min on crude cell-free lysate produced by sonication. (A) In SDS-PAGE the MW marker (lane 1) enabled identification of His_6-tagged *M. tuberculosis* DHDPS (34 kDa). Lanes 2 to 6 were loaded with supernatant (1 in 20 dilutions) from centrifugation at 4 °C and speeds as follows: 2 - 1000 rpm, 3 - 2000 rpm, 4 - 3000 rpm, 5 - 4000 rpm, 6 - 5000 rpm. Lanes 7 to 10 were loaded with pellet (re-suspended in 500 µL of high urea, high SDS) from the same centrifugation at the corresponding speeds: 7 - 2000 rpm, 8 - 3000 rpm, 9 - 4000 rpm, 10 - 5000 rpm (B) DHDPS activity in the supernatant decreased with increasing speeds of centrifugation, as determined by the *o*-aminobenzaldehyde assay. Purified *E. coli* and *M. tuberculosis* DHDPS were used as positive controls.[1,30]**

The loss of *M. tuberculosis* DHDPS with centrifugation corresponds to the increasing sedimentation of insoluble proteins, and thus it seems to reflect the insoluble nature of the recombinantly expressed protein. Filtration was considered as an alternative to centrifugation,

but the initial results were not promising and the higher concentration of insoluble proteins in samples centrifuged for 5 minutes at 5000 rpm resulted in problems during chromatography; hence, further efforts focused on improving the yield of soluble *M. tuberculosis* DHDPS.

Co-expression of chaperonins GroES and GroEL

The most significant improvement in the yield of soluble protein came from the co-expression of GroES and GroEL proteins**.** These proteins are part of a subclass of proteins called chaperonins, which bind non-native forms of proteins and facilitate their refolding into the native conformation.[36] Lorimer and colleagues designed a plasmid containing the *groEL* and *groES* genes, under the control of the p15A replicon, which is compatible for co-expression with the pET and pETM vector series.[37] The plasmid pGroESL was acquired from Dr Celia Webby, who found that co-expression of GroEL and GroES chaperonins produced a small amount of soluble, biologically active *M. tuberculosis* 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase (DAH7PS), which otherwise was insoluble.[38]

The plasmid pGroESL was used to transform *E. coli* BL21 (DE3) competent cells (Novagen). Transformants were selected using LB agar plates containing chloramphenicol. Subsequently, BL21 (DE3) pGroESL competent cells were transformed by electroporation (Dr Celia Webby, *pers. comm.*) using the plasmid pMTB02 **(Appendix A)**, conferring kanamycin-resistance. Transformed cells containing both pGroESL and pMTB02 were selected for using LB agar plates containing chloramphenicol and kanamycin, and glycerol stocks were made.

Two other *E. coli* strains were transformed using pMTB02 for comparison of soluble protein expression: *E. coli* BL21-CodonPlus® (DE3) RP and BL21 (DE3) pLysS, both of which have a chloramphenicol-resistance phenotype. *E. coli* BL21-CodonPlus® (DE3) RP cells, are engineered to contain extra copies of genes for tRNAs that frequently limit translation of recombinant proteins in *E. coli*,[39] and were used in the original purification method developed by Kefala.[1] *E. coli* BL21 (DE3) pLysS cells have the benefit of accumulating low levels of T7 lysozyme, which prevents low level expression of the cloned gene product before induction,[40] and facilitates the production of cell extracts.[41,42] The plasmid pMTB02 was used to transform BL21-CodonPlus® (DE3) RP competent cells (Stratagene), and BL21 (DE3) pLysS competent cells (prepared using calcium chloride), by the heat shock method. As outlined in the previous paragraph, chloramphenicol and kanamycin were used for selection,

and glycerol stocks were prepared. The production of soluble, biologically active protein was examined in each of these different bacterial strains by both SDS-PAGE **(Figure 2.4)** and the coupled assay **(Table 2.1)**.

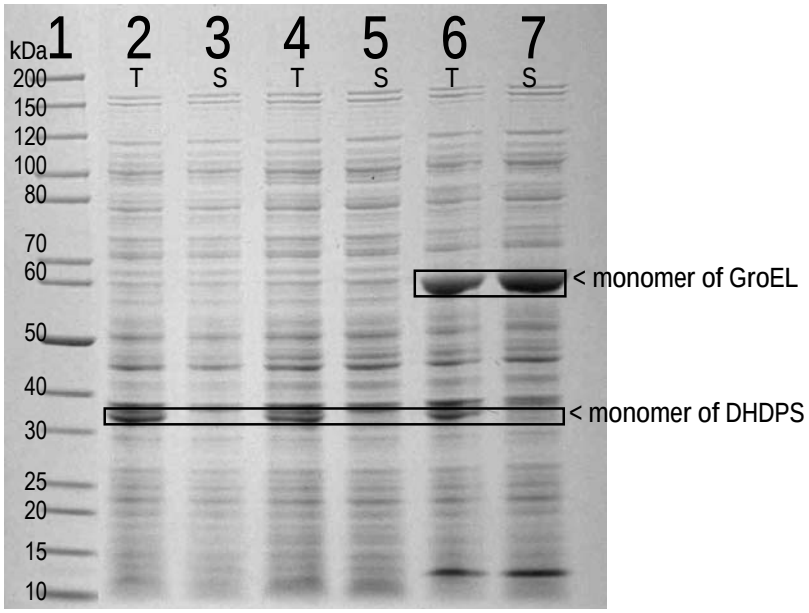


Figure 2.4: Comparison of soluble *M. tuberculosis* DHDPS expression by SDS-PAGE analysis of the total (T) and soluble (S) fraction from the cell-free lysate produced by sonication from different *E. coli* strains, loaded as follows: BL21-CodonPlus® (DE3) RP (lanes 2, 3), BL21 (DE3) pLysS (lanes 4, 5) and BL21 (DE3) pGroESL (lanes 6, 7). The MW marker (lane 1) allows for identification of His_6-tagged *M. tuberculosis* DHDPS (34 kDa). Cultures of each strain were grown for 20 hours at room temperature following induction. The soluble (S) fraction resulted from centrifugation for 30 min, at 11000 rpm and 4 °C.

BL21 (DE3) pGroESL was shown to be the best *E. coli* strain for expressing soluble, biologically active *M. tuberculosis* DHDPS. Unlike the other bacterial strains, BL21 (DE3) pGroESL had an observable band corresponding to DHDPS from *M. tuberculosis* in the soluble fraction **(Figure 2.4, lane 7)**. Most significantly, the activity per gram of cells harvested for *E. coli* BL21 (DE3) pGroESL was much greater than that for BL21-CodonPlus® (DE3) RP **(Table 2.1)**, even though the over-expression level as judged by SDS-PAGE **(Figure 2.4)** was very similar.

Table 2.1: Comparison of enzyme activity from different bacterial strains.

E. coli strain	Harvested cells[a] (g)	Total activity[b] (units)[c] in crude[d]	Total activity[b] (units)[c] in soluble fraction[e]
BL21-CodonPlus® (DE3) RP	5.23	2.6	0.77
BL21 (DE3) pLysS	6.51	6.1	5.4
BL21 (DE3) pGroESL	5.30	9.0	8.5

[a] Wet-weight of cells from 1 L of culture grown in LB media for 33 hours at room temperature after induction.

[b] Enzyme activity was determined using the quantitative coupled assay, detailed in section 2.2.3.

[c] 1 unit is defined as the consumption of 1 µmol of NADPH per second.

[d] Crude preparation required cells to be washed and re-suspended in buffer A to give a total volume of 50 mL, followed by sonication for 3 minutes in 2 s pulses with 10 s delays at 4 °C.

[e] Soluble fraction was the supernatant from centrifugation for 30 min at 11000 rpm and 4 °C.

Buffer components

The purification buffers used in the original protocol contained several components, in addition to the buffer, which were intended to improve solubility and protect against enzyme degradation. The effect of these components on thermal stability was investigated using differential scanning fluorescence (DSF), where a sharp increase in fluorescence **(**as shown in **Figure 2.5A)** corresponds to the protein going from a folded to unfolded state, and therefore the mid-point of this transition reflects the apparent melting temperature (T_m^{app}).[43] The denaturation of *M. tuberculosis* DHDPS in 20 mM Tris.HCl, pH 8.0 **(Figure 2.5A)**, gave a T_m^{app} of 71.5 (±0.1) °C. The inclusion of 250 mM NaCl in the buffer increased T_m^{app} by ~6 °C **(Figure 2.5B)**, which is unsurprising given that the majority of proteins are more stable at physiological ionic strength,[23] and confirmed that NaCl should be included in the buffer. The subsequent additions of glycerol and β-ME (giving final concentrations of 5 % and 2 mM, respectively) had minimal and counteracting effects **(Figure 2.5B)**. However, both continued to be included in the purification buffers, since preparations stored in glycerol were observed to be appreciably less degraded over time and β-ME protects against oxidative effects.[44]

A modification to the original purification buffer was considered; that is the inclusion of one of the enzyme's substrate, pyruvate, as it had been found to increased protein yield and stability in previous studies with DHDPS.[19] Interestingly, *M. tuberculosis* DHDPS gave a T_m^{app} of 85.3 (±0.1) °C in 10 mM pyruvate, 2 mM β-ME, 5 % glycerol, 250 mM NaCl, 20 mM Tris.HCl, pH 8.0, increasing the T_m^{app} by ~8 °C in comparison to that determined in

the same buffer without pyruvate **(**77.5 (±0.3) °C in 2 mM β-ME, 5 % glycerol, 250 mM NaCl, 20 mM Tris.HCl, pH 8.0, as shown in **Figure 2.5B)**, which indicated that pyruvate had a substantial stabilizing effect. Consequently, the modified protocol included 10 mM pyruvate in all purification buffers, since its stabilizing effect is likely to improve the yield of soluble, non-denatured *M. tuberculosis* DHDPS. Although, it was excluded from the final storage buffer as a background concentration of pyruvate would complicate kinetic characterization.

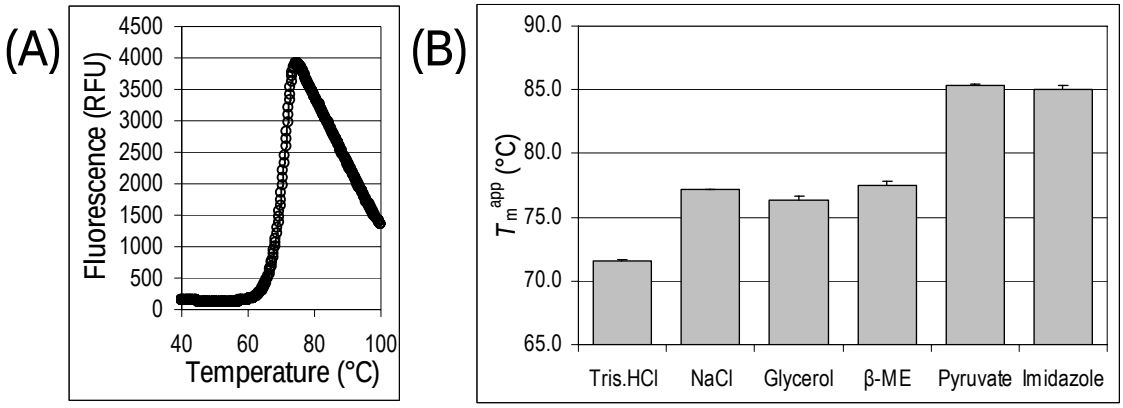


Figure 2.5: Determination of T_m^{app} using differential scanning fluorescence. (A) The melting curve of 0.5 mg.mL^{-1} *M. tuberculosis* DHDPS (His$_6$-tag removed) in the presence of 20 mM Tris.HCl, pH 8.0. (B) The additive effect of different buffer components on T_m^{app} for *M. tuberculosis* DHDPS. Each T_m^{app} was measured in triplicate and the error bars show standard deviation. Beginning with 20 mM Tris.HCl, pH 8.0 (leftmost bar), subsequent buffers had the indicated component added to the previous buffer systems. Thus the T_m^{app} of 85.1 (±0.2) °C on the far right corresponds to an extraction buffer composed of 20 mM Tris.HC1, 250 mM NaCl, 5 % glycerol, 2 mM β-ME, 10 mM pyruvate and 10 mM imidazole.

The components of the extraction buffer were also considered, as the preparation of cell extracts influences both the total quantity of protein recovered and the biological activity retained.[45] The extraction buffer included a low concentration (10 mM) of imidazole to minimize non-specific binding to the chromatography resin in the subsequent step, and thus improved enzyme purity.[46] Since the effect of imidazole on the T_m^{app} observed with DSF was insignificant **(Figure 2.5B)**, it continued to be added in the extraction buffer. The original purification method also included the commercially available mild detergent, BugBuster™ (Novagen);[1] however, no significant improvements in extraction of active *M. tuberculosis* DHDPS were observed (Dr Simone Weyand, *pers. comm.*), and therefore the detergent was

excluded from the purification procedure, as was the EDTA-free protease inhibitor (Roche), since no degradation/proteolysis was observed.

2.3.3 Modified purification procedure for *M. tuberculosis* DHDPS

The most important modification to the original method was the co-expression of refolding chaperonins GroES and GroEL with *M. tuberculosis* DHDPS. As in the original protocol, over-expression of DHDPS was induced by IPTG and cells were grown overnight at 20 °C, and then harvested by centrifugation. Crude cell-free lysates were produced by 5 minutes of ultrasonication at 4 °C, in 20 s pulses **(Figure 2.6, lane 2)**, in extraction buffer (buffer A) including 10 mM pyruvate (as described in the previous section), followed by clarification with centrifugation at 10000 rpm. In contrast to the original protocol, soluble *M. tuberculosis* DHDPS could be observed in the supernatant by SDS-PAGE **(Figure 2.6, lane 3)**.

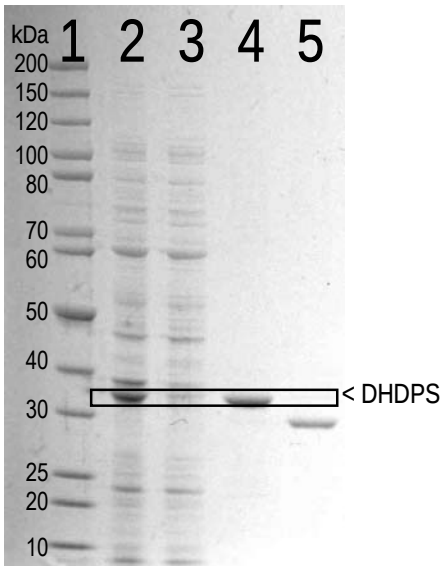


Figure 2.6: **A typical purification using the modified protocol for *M. tuberculosis* DHDPS, over-expressed in *E. coli* BL21 (DE3) pGroESL pMTB02, examined using SDS-PAGE. The MW marker, in lane 1, enabled identification of His_6-tagged *M. tuberculosis* DHDPS as the 34 kDa band (black rectangle). Protein preparations from various stages of purification were loaded into other lanes as follows: 2 - crude cell-free extract, 3 - after centrifugation, 4 - eluted from His_6-tag affinity column, 5 - after overnight incubation with TEV protease. In lane 5, the cleavage of the 4 kDa His_6-tag gives rise to a 30 kDa band corresponding to non-tagged *M. tuberculosis* DHDPS.**

The supernatant from centrifugation was loaded onto a His_6-tag affinity column (HisTrap Crude, GE Healthcare) pre-equilibrated with buffer A. The column was washed with three buffers at pH 8.0 (chapter six, section 6.4), and buffer D eluted His_6-tagged *M. tuberculosis* DHDPS by increasing imidazole concentration to 250 mM, and resulted in 22-fold increase in enzyme purity **(Table 2.2,** and **Figure 2.6, lane 4)**. The yield of ~5 mg of His_6-tagged protein from 1 L of culture **(Table 2.2)** was considerably greater than the 1.6 mg obtained using the original procedure (section 2.2.1).

Table 2.2: Optimized purification of *M. tuberculosis* DHDPS from *E. coli* BL21 (DE3) pGroESL pMTB02 (4.5 L culture, 9.2 g of cells).

Purification step	Volume (mL)	Protein[a] (mg)	Total activity[b] (units[c])	Specific activity (units[c]/mg)	Relative total activity[d]	Purification[d]
Crude	150	990	97.5	0.1	-	1-fold
Centrifuged	110	575	44.9	0.08	46%	1-fold
Affinity (His_6-tagged)	25.0	24	52.4	2.2	54%	22-fold
Affinity (cleaved)	25.0	21	82.2	3.9	84%	39-fold
Concentrated	3.0	2.2	6.9	3.2	7%	32-fold
Gel filtration	11.0	3.3	11.8	3.6	12%	36-fold

[a] Protein concentration was determined using the Bradford assay, detailed in chapter six, section 6.3.1.

[b] Enzyme activity was determined using the quantitative coupled assay, detailed in section 2.2.3.

[c] 1 unit is defined as the consumption of 1 µmol of NADPH per second.

[d] Crude preparation may contain inhibitors of NADPH consumption/NADPH-utilizing enzymes, both of which interfere with the coupled assay; therefore, values in the last two columns may be inaccurate due to their basis on crude material.

Active fractions were detected using the *o*-aminobenzaldehyde assay (section 2.2.1) and buffer exchanged back into buffer A using desalting columns. The His_6-tag was cleaved off by overnight incubation with recombinant TEV protease **(Appendix B)**, in the presence of 0.5 mM EDTA and 1 mM DTT, leaving four linker amino acids, G-A-M-A, at the N-terminus.[35,47] After SDS-PAGE confirmed the cleavage of the tag **(Figure 2.6, lane 5)**, the cleaved *M. tuberculosis* DHDPS was passed through a His_6-tag affinity column, pre-equilibrated with buffer A, to remove the His_6-tag peptide fragment. Cleaved non-tagged *M. tuberculosis* DHDPS had higher specific activity, as reflected by the increase in specific activity between purification steps 3 and 4 **(Table 2.2)**.

Unfortunately, the protein required a significant concentration step in preparation for gel filtration and a substantial loss of yield resulted after the ultrafiltration spin column (Vivaspin 15, Vivascience) **(Table 2.2)**. This was likely caused by proteins binding to the filtration membrane, as no obvious precipitate formed, and no enzyme activity was found in the flow through. The concentrate was loaded onto a gel-filtration column (Superdex 200, 16/60, GE Healthcare) to remove any remaining impurities; the active fraction, which eluted with molecular mass of ~120 kDa, consistent with a homotetramer, was collected. The purified *M. tuberculosis* DHDPS was analyzed by SDS-PAGE and stored in storage buffer [20 mM Tris.HCl, pH 8.0, 250 mM NaCl, 5 % glycerol and 2 mM β-ME] at 4 or -20 °C for several weeks without detectable degradation. The final yield of ~0.7 mg of non-tagged protein per L of LB media from the optimized procedure was comparable to the reported yield from the original procedure,[1] and enabled preparation of milligram quantities of soluble, folded, active *M. tuberculosis* DHDPS for biochemical and biophysical characterization.

2.3.4 Over-expression & purification of DHDPR

DHDPR, the coupling enzyme for steady-state kinetic studies of DHDPS, was obtained *via* over-expression from the plasmid pTM1521. This plasmid carries the gene for *T. maritima* DHDPR with a non-cleavable His_6-tag at its N-terminus, and was kindly donated by Scott Lesley and Heath Klock (Joint Center for Structural Genomics, Genomics Institute of the Novartis Research Foundation, San Diego, USA). Since the enzyme was used in excess, the species from which the DHDPR was derived was unimportant. *T. maritima* DHDPR was chosen for convenience and its thermal stability, which allowed the addition of a heat shock step to the published protocol.[48] Incubation for 2 minutes at 95 °C resulted in no significant loss of DHDPR activity and protein purity was improved several folds **(Table 2.3)**. Highly homogenous DHDPR was eluted from the affinity column, as judged by denaturing SDS-PAGE, and stored at -20 °C.

2.3.5 Over-expression & purification of *E. coli* DHDPS

For comparison, *E. coli* DHDPS was also produced, using a procedure modified from the protocol developed by Coulter *et al*.[4] The enzyme was expressed and purified from XL1-Blue cells containing pJG001, which contains the *dap*A gene from *E. coli*. In an analogous fashion to DHDPR, the relative thermal stability of *E. coli* DHDPS allowed for inclusion of a heat shock step (for 2 minutes at 70 °C) with no loss of specific activity **(Table 2.4)**. The

supernatant was loaded onto ion exchange column (Q-Sepharose, GE Healthcare), and eluted with a linear gradient from 0 to 1 M NaCl. The active peak, determined by the *o*-aminobenzaldehyde assay (section 2.2.1), was collected and dialyzed overnight in buffer containing 0.5 M ammonium sulfate, in preparation for hydrophobic interaction chromatography (phenyl Sepharose, GE Healthcare). Any remaining contaminants were removed by the second chromatography step and the active fractions, containing *E. coli* DHDPS, were pooled and stored at -20 °C.

Table 2.3: Typical purification of *T. maritima* DHDPR (4.5 L culture, 14.0 g of cells).

Purification step	Protein[a] (mg)	Total activity[b] (units[c])	Specific activity (units/mg)	Relative total activity[d]	Purification[d]
Crude	321	6.39	0.020	-	1-fold
Heat shock	129	6.65	0.051	104%	3-fold
Affinity	42	5.08	0.12	79%	6-fold

Table 2.4: Purification of *E. coli* DHDPS (4.5 L culture, 10.7 g of cells).

Purification step	Protein[a] (mg)	Total activity[b] (units[c])	Specific activity (units/mg)	Relative total activity[d]	Purification[d]
Crude	855	324	0.38	-	1-fold
Heat Shock	361	203	0.56	63%	1-fold
Ion Exchange	113	147	1.30	45%	3-fold
Hydrophobic Exchange	74	104	1.41	32%	4-fold

[a] Protein concentration was determined using the Bradford assay, detailed in chapter six, section 6.3.1.

[b] Enzyme activity was determined using the quantitative coupled assay, detailed in section 2.2.3.

[c] 1 unit is defined as the consumption of 1 µmol of NADPH per second.

[d] Crude preparation may contain inhibitors of NADPH consumption or NADPH-utilizing enzymes, both of which interfere with the coupled assay; therefore, values in the last two columns may be inaccurate due to their basis on crude material.

2.4 Biochemical analysis of *M. tuberculosis* DHDPS

The determination of the steady-state kinetics of *M. tuberculosis* DHDPS required a method for measuring the initial-rate of reaction, such as the coupled assay (section 2.2.3), as well as an understanding of the factors that affect this rate, such as pH, ionic strength and temperature. This provided a comprehensive basis of comparison for subsequent work. The effect of (*S*)-lysine and other aspartic family amino acids on enzyme activity was also

investigated; both in light of the putative regulatory role of DHDPS in the DAP pathway and the (*S*)-lysine binding pocket providing a possible binding site for inhibitors (section 2.1).

2.4.1 Optimum pH

Most enzymes are active within a limited pH range and many have optimal activity at a specific pH, because changes in pH can influence the enzyme structure and stability, and its affinity for its substrate.[22] A series of buffers covering a pH range of 6 to 9 with the same ionic strength (adjusted by the addition of NaCl)[49] was used to determine the optimum pH at 30 °C in the coupled assay system. The activity of *M. tuberculosis* DHDPS had an optimum pH between 7.75 to 8.25 **(Figure 2.7A)**, which is similar to the literature value for *E. coli* DHDPS of 8.4.[6] Therefore, there seemed no reason to modify the coupled assay procedure used for steady-state kinetics, where the pH is maintained at 8.0 by 100 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethane sulfonic acid (HEPES) buffer.[30]

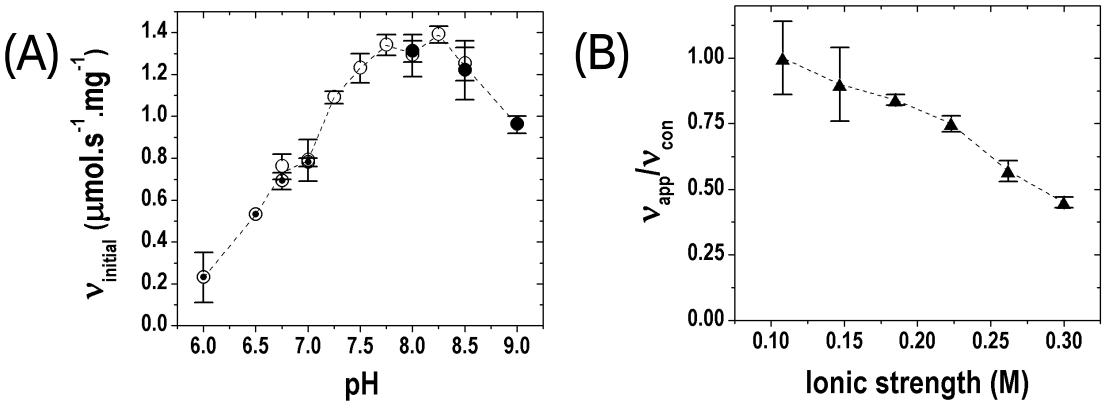


Figure 2.7: **The optimum conditions for enzyme activity determined by varying (A) pH with different buffers (◉ MES, ○ HEPES, ● Bicine), and (B) ionic strength with salt concentration. The data in the second plot are normalized by dividing the apparent rate (v_{app}) by the highest measured rate (v_{con}). Each data point was measured at least in duplicate and the error bars show standard deviation. The dashed line shows the general trend of the data.**

2.4.2 The effect of ionic strength

The ionic strength inside the cytoplasm of a typical cell ranges from 0.15-0.2 M.[23] Below physiological ionic strength, enzyme stability, but not activity, tends to be affected; however, above 0.2 M, enzyme activity is often depressed.[23] The effect of ionic strength on the

enzymatic reaction was determined using the coupled assay system and a series of 20 mM HEPES buffers covering a range of ionic strengths (adjusted by NaCl and corrected for the amount of HCl/NaOH added to bring the pH to 8.25 at 30 °C). Greater variability (±14 %) was seen in initial-rate measurements, performed in quadruplicate, at ionic strengths below 0.20 M, as compared to the variability (±5 % or less) in triplicate measurements at or above 0.20 M **(Figure 2.7B)**, possibly reflecting enzymatic destabilization at low ionic strengths.

The optimal ionic strength for stabilization and activity seems to be in the physiological range of 0.15-0.20 M, since below this ionic strength the assays are less reproducible and above it the activity of *M. tuberculosis* DHDPS decreases. As a consequence, 50 mM of NaCl was included in the 200 mM HEPES buffer stock to increase the ionic strength in the coupled assay system for kinetic characterization **(**details in **Appendix D)**.

2.4.3 Thermal stability

Thermal degradation was investigated by incubating DHDPS for a defined period at various temperatures before measuring initial-rates at the standard assay temperature of 30 °C. *M. tuberculosis* DHDPS showed a greater thermal stability than *E. coli* DHDPS, not degrading until 80 (±1) °C **(Figure 2.8A)**.

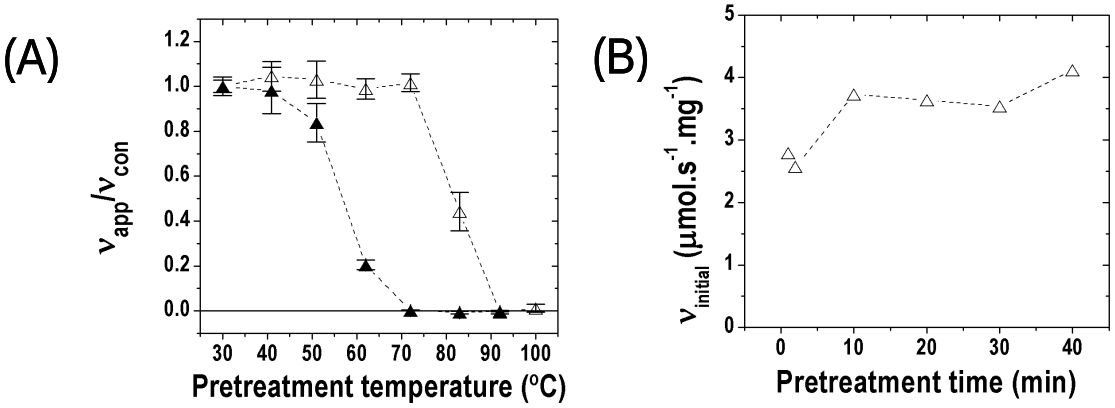


Figure 2.8: **The thermal stability of *M. tuberculosis* DHDPS (△) was compared to that of *E. coli* DHDPS (▲) using the apparent initial-rate (v_{app}) divided by rate of enzyme activity without pre-treatment (v_{con}). The enzyme activity was determined with the coupled assay (A) after pre-treatment by 5 minute incubation at increasing temperature and (B) after incubation at 70 °C for increasing durations. The dashed line shows the general trend of the data.**

This result is in good agreement with the T_m^{app} of 77.5 (±0.3) °C determined using DSF for *M. tuberculosis* DHDPS in storage buffer [20 mM Tris.HCl, 250 mM NaCl, 5 % glycerol, 2 mM β-ME] (section 2.3.3, Figure 2.5). Increasing the temperature to 80 °C, will have changed the pH buffered by HEPES during incubation from 8.25 to pH 7.5,[49] which is still sufficiently near optimum to disregard any pH effects. Further experiments with incubation at 70 °C (~pH 7.7) showed that *M. tuberculosis* DHDPS maintained its activity for at least 40 minutes **(Figure 2.8B)**. The slight increase in activity observed is likely due to evaporation concentrating the protein sample over the course of the incubation.

2.4.4 Steady-state kinetics of wild-type *M. tuberculosis* DHDPS

Steady-state kinetics enable simple calculation of kinetic parameters *in vitro* and inform our understanding of enzyme action and metabolism.[50] Enzymatic rates are measured just subsequent to assay initiation, before the rate is affected by product accumulation or substrate consumption. There is only a relatively short interval, during which the assumptions of kinetic models, such as the Michaelis-Menten, are valid.[49] Consequently, prompt measurement and a recordable period of linearity is required for accurate initial-rates and kinetic analysis.

Kinetic models describe the binding of substrates and release of products in mathematical terms, such as the Michaelis-Menten equation **(Equation 2.1)**, which introduced the Michaelis constant (K_M) and the maximum velocity attainable (V) to describe the relationship between the initial-rate or steady-state velocity (v), substrate concentration ([S]):[50]

$$v = \frac{V}{K_M + [S]}$$ **Equation 2.1** [24]

Michaelis and Menten developed this equation to reflect a single substrate (S) binding to the enzyme (E) to form an enzyme-substrate complex (E:S), in rapid reversible equilibrium, with subsequent slow conversion and release of product (P) **(Figure 2.9)**:[24]

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} E{:}S \xrightarrow{k_2} E + P$$

Figure 2.9: A schematic diagram of the kinetic mechanism described by the Michaelis-Menten equation, with the rate constant of the second step, k_2, is much smaller than k_1.[24]

In this case **(Figure 2.9)**, where $k_2 << k_1$ and therefore the second step determines the rate of the overall reaction, V is equal to k_2 and K_M is the dissociation constant (k_1/k_{-1}) of the enzyme-substrate complex.[24] Both K_M and V are used to describe much more complex kinetic models than Michaelis-Menten; however, their meaning is no longer assigned to a particular dissociation or reaction step, respectively.[50] For example the ping-pong model involves two substrates (A, B) binding in a specific order, with a release of a product to form an enzyme-substrate intermediate (F), prior to the binding of the second substrate (B) **(Figure 2.10)**, and therefore V and K_M are considered in the more general terms of the overall process.[51]

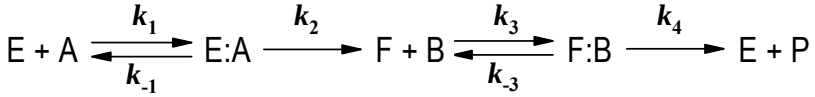


Figure 2.10: A schematic diagram of the ping-pong mechanism.

Kinetic data for *M. tuberculosis* DHDPS were fitted to the ping-pong (or substituted-enzyme) model **(Equation 2.2)**, the compulsory-order ternary-complex model **(Equation 2.3)** and the ping-pong with substrate inhibition model **(Equation 2.4)**. These models all describe two substrate systems where substrates bind in a specific order; however, the order of binding is ambivalent for the first model due to the symmetry of the equation. In the ternary-complex model, the second substrate (B) binds to the enzyme-substrate complex (E:A) prior to the release of any product.[24]

$$v = \frac{V \cdot [A] \cdot [B]}{K_M^B \cdot [A] + K_M^A \cdot [B] + [A] \cdot [B]} \qquad \textbf{Equation 2.2}\ ^{24}$$

$$v = \frac{V \cdot [A] \cdot [B]}{K_M^B \cdot [A] + K_M^A \cdot [B] + [A] \cdot [B] + K_s^A \cdot K_M^B} \qquad \textbf{Equation 2.3}\ ^{24}$$

$$v = \frac{V \cdot [A] \cdot [B]}{K_M^B \cdot [A] + K_M^A \cdot [B] \cdot (1 + [B]/K_i^B) + [A] \cdot [B]} \qquad \textbf{Equation 2.4}\ ^{24}$$

Here v is described in terms of two variable substrate concentrations [A] and [B], corresponding to pyruvate and (*S*)-ASA for DHDPS, and by at least three constants, V and K_M^A and K_M^B, the Michaelis constants for the respective two substrates. More specifically, K_M^A corresponds to the Michaelis constant when B is saturating, and *vice versa* for K_M^B, whereas K_s^A, which only appears in **Equation 2.3**, corresponds to this constant for substrate A in the

absence of substrate B.[24] K_i^B, which appears in **Equation 2.4**, is a dissociation constant that defines the strength of inhibition caused by the second substrate binding first and forming a dead-end complex.[24]

Initial-rate data for *M. tuberculosis* DHDPS were determined with the coupled assay, using the buffer stock described at the end of section 2.4.2, for a five-by-five matrix varying pyruvate and (*S*)-ASA concentrations.[26] As recommended by Cornish-Bowden, substrate concentrations were chosen based on apparent K_M values determined by varying one substrate while the other was held constant.[24] Accuracy is best when the range of concentrations corresponds to the greatest variation in rate, so the concentration of (*S*)-ASA and pyruvate was varied between 0.2 to 10 times the apparent K_M (which was determined to be 0.3 mM for both substrates).[24] The initial-rate of product formation was measured in triplicate, and was reproducible (±10 %). The data were analyzed by non-linear curve fitting to appropriate kinetic models using the program ENZFITTER (Biosoft, Cambridge, U.K.).

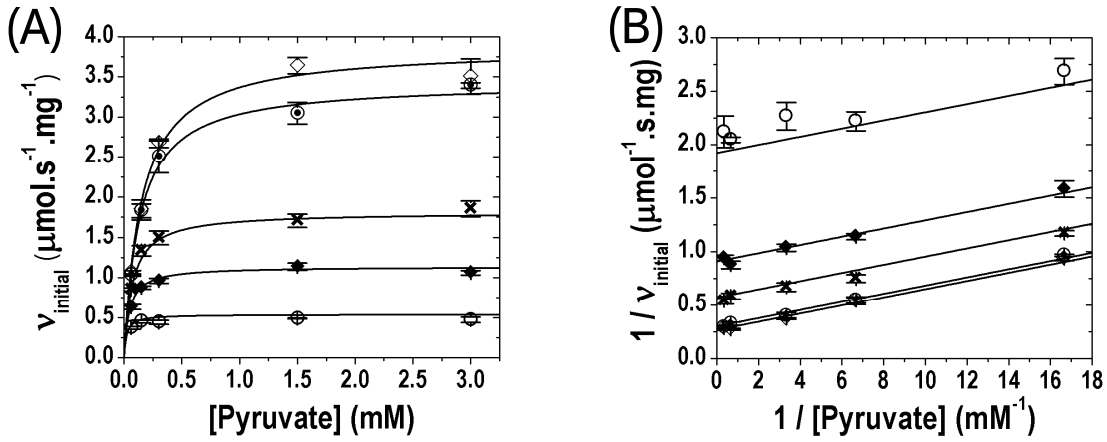


Figure 2.11: The initial velocity of wild-type *M. tuberculosis* DHDPS was measured with the coupled assay over varying concentrations of both substrates, pyruvate and (*S*)-ASA, shown on the x-axis and with symbols (◇ 3.0 mM, ⦿ 1.5 mM, × 0.30 mM, ◆ 0.15 mM, ○ 0.06 mM (*S*)-ASA). Each point was measured at least in triplicate and the data were fitted with the ping-pong model giving an R^2 of 0.99 using the program ENZFITTER. Plot (B) shows the Lineweaver-Burk transformation, fitted with the parallel lines predicted for the ping-pong model.

The ping-pong model provided the best fit **(Figure 2.11A** and **2.12A)**, shown by its R^2 value of 0.99, and yielded a V of 4.2 (±0.1)[a] µmol.s^{-1}.mg^{-1} and K_M constants of 0.16 (±0.01) mM and 0.43 (±0.03) mM for pyruvate and (*S*)-ASA, respectively. These are similar in magnitude

to the K_M constants determined for other DHDPS enzymes by the coupled assay system (see Table 1.2, chapter one, section 1.3.1). The substrate inhibition model also gave a reasonable fit, but the $K_i^{(S)\text{-}ASA}$ value of 25 mM is outside the concentration range examined in this study.

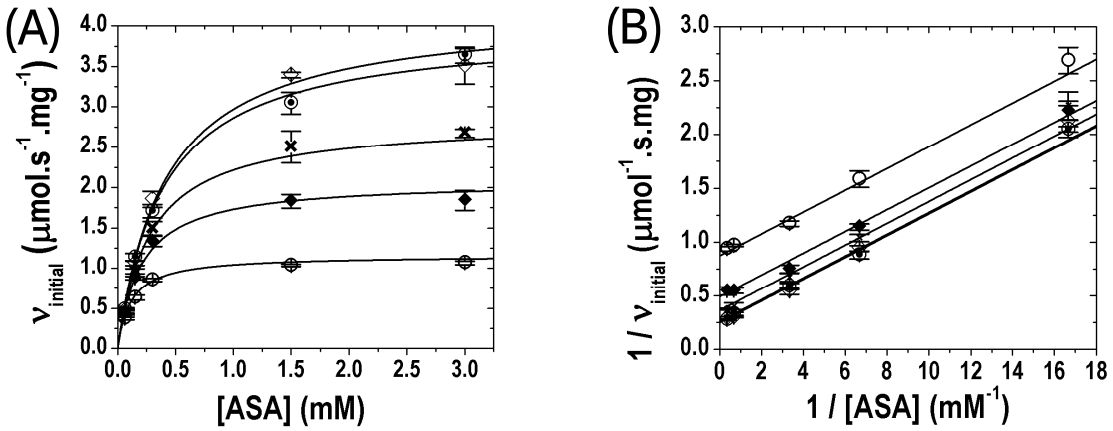


Figure 2.12: The initial velocity of wild-type *M. tuberculosis* DHDPS was measured over varying concentrations of both substrates, (*S*)-ASA and pyruvate, shown on the x-axis and with symbols (◇ 3.0 mM, ◉ 1.5 mM, × 0.30 mM, ◆ 0.15 mM, ○ 0.06 mM pyruvate). Each point was measured in triplicate and the data were fitted with the ping-pong model using ENZFITTER. Plot (B) shows the Lineweaver-Burk transformation, fitted with the parallel lines predicted for this model.

The molecular weight (expressed at enzyme molarity, e_0) of *M. tuberculosis* DHDPS can be used to convert V into the catalytic turnover number, k_{cat}, defined as the number of catalytic processes or substrates "turned-over" by an enzyme per unit time.[24] *M. tuberculosis* DHDPS has a k_{cat} of 132 (±3)* s^{-1} per active site in comparison with *E. coli* DHDPS k_{cat} of 124 (±7) s^{-1} **(Table 2.5)**.[28] The efficiency of any enzyme is shown by k_{cat}/K_M. The efficiency of a perfect enzyme cannot be any faster than the rate at which substrates and products diffuse in and out of the active site, thus the upper limit due to diffusion means there is a maximum value possible for k_{cat}/K_M of 10^8 to 10^9 $s^{-1}.M^{-1}$.[50] Dividing k_{cat} by K_M for pyruvate and for (*S*)-ASA gave values for catalytic efficiency with magnitudes of 10^5 $s^{-1}.M^{-1}$ showing that *M. tuberculosis* DHDPS, like many enzymes, is not at the diffusion limit.

* The V and k_{cat} reported here were calculated using initial-rates determined from recorded data by a linear fit to first 10 s measured. Subsequent work (in chapter 5) utilized this approach. The values published elsewhere were calculated using initial rates determined from the same data set by fitting an algorithm from ENZFITTER called initial rate for the first 50 s, which resulted in a V of 4.42(±0.1) µmol.s^{-1}.mg^{-1}.

Table 2.5: Kinetic parameters determined by fitting initial-rate measurements, compared with literature values for *E. coli* DHDPS.[28]

	M. tuberculosis DHDPS	*E. coli* DHDPS
Kinetic model	Ping-pong	Ping-pong
K_M for (*S*)-ASA (mM)	0.43 (±0.02)	0.11 (±0.02)
K_M for pyruvate (mM)	0.17 (±0.01)	0.26 (±0.01)
Monomer molecular weight (Da)	31156.5	31269.0
e_0 (mgs of subunit per µmol)	31.16	31.27
k_{cat} per active site (s^{-1})	132 (±3)[a]	124 (±2)
k_{cat}/K_M for (*S*)-ASA ($s^{-1}.M^{-1}$)	3.0 (±0.2) $\times 10^5$	1.1 (±0.2) $\times 10^6$
k_{cat}/K_M for pyruvate ($s^{-1}.M^{-1}$)	7.8 (±0.6) $\times 10^5$	4.8 (±0.6) $\times 10^5$

Further confirmation that substrates bind to *M. tuberculosis* DHDPS in a ping-pong fashion was provided by Lineweaver-Burk **(Figure 2.11B** and **2.12B)** and Hanes plots **(Figure 2.13)**. These graphical methods aid in visual detection of the characteristic trends that reflect the different kinetic models and any deviations from the model fitted.[24,52] The family of hyperbolae is transformed into a series of straight lines, most commonly by plotting the reciprocal of v and [S], utilizing the method developed by Lineweaver and Burk.[24,49,52] This approach compresses high substrate concentration data into a small region, while low concentration data, with the slowest rates and therefore highest errors, are emphasized.[24,49,52] Experimental error is much less distorted in the Hanes plot ([S]/v against [S]); thus it is considered more accurate and generally superior.[24,49,52]

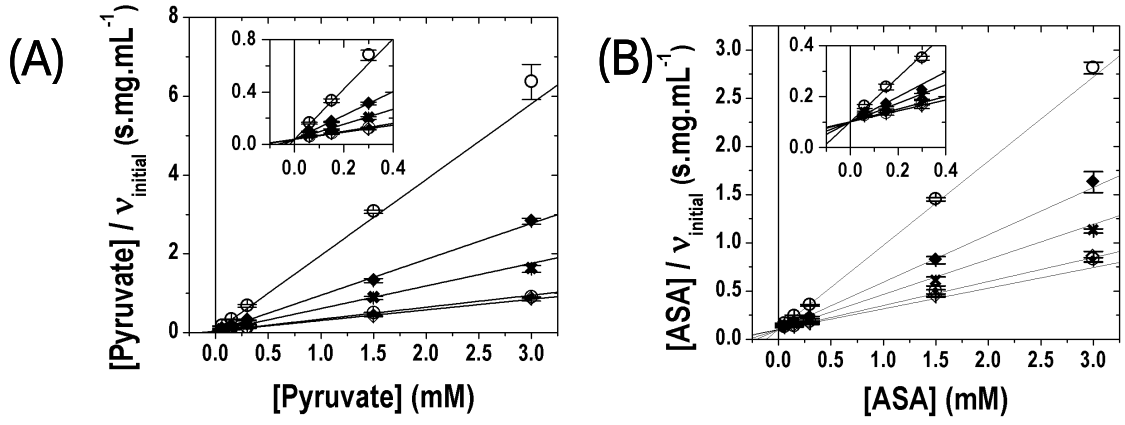


Figure 2.13: The Hanes transformation, reflecting the trend of y-axis intersection predicted for the ping-pong model, plotted against (A) varying concentrations of (*S*)-ASA on the x-axis (◇ 3.0 mM, ◉ 1.5 mM, ✕ 0.30 mM, ◆ 0.15 mM, ○ 0.06 mM pyruvate) and (B) varying concentrations of pyruvate on the x-axis (◇ 3.0 mM, ◉ 1.5 mM, ✕ 0.30 mM, ◆ 0.15 mM, ○ 0.06 mM (*S*)-ASA).

The ping-pong model manifests itself by parallel lines in the Lineweaver-Burk plot **(Figure 2.11B** and **2.12B)** and a set of lines intersecting on the y-axis in the Hanes plot **(Figure 2.13)**; the lack of curvature seen in both plots indicates the absence of substrate inhibition.[24,52] The ping-pong mechanism of *M. tuberculosis* DHDPS is consistent with the kinetic mechanism proposed for other DHDPS enzymes, discussed in chapter one, section 1.3.1 & 1.3.2.

2.4.5 The effect of (*S*)-lysine and other amino acids on *M. tuberculosis* DHDPS

The role of DHDPS as a control point in DAP/(*S*)-lysine biosynthesis in mycobacteria has not been elucidated; however, differences in the regulation of this pathway as compared to *E. coli* have been emphasized by gene-knockout experiments (chapter one, section 1.2.3).[53] Gram-positive bacteria tend to have DHDPS that are uninhibited by (*S*)-lysine (chapter one, section 1.2.4), and traditionally mycobacteria, such as *M. tuberculosis*, have been grouped with Gram-positive bacteria, although, this classification is debated.[54]

The coupled assay was performed in the presence of increasing amounts of (*S*)-lysine, with pyruvate and (*S*)-ASA at concentrations of 0.3 and 0.2 mM, respectively. Except at very high concentrations of (*S*)-lysine (50 mM and above), (*S*)-lysine did not have any effect on the activity of *M. tuberculosis* DHDPS. The IC_{50}, that is the concentration needed for 50 % inhibition, was determined to be ~250 mM **(Figure 2.14)**, which greatly exceeds the level of (*S*)-lysine that could be reasonably expected in a cell. Thus, *M. tuberculosis* DHDPS can be considered insensitive to (*S*)-lysine, fitting with the trend found in Gram-positive bacteria.

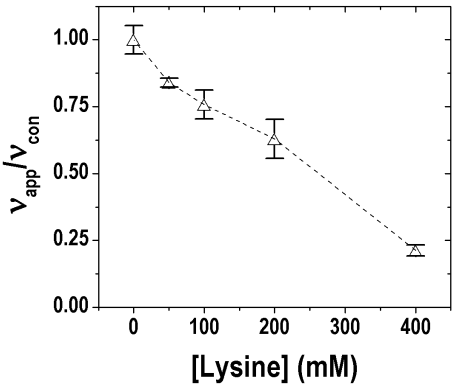


Figure 2.14: The activity of *M. tuberculosis* DHDPS, in the presence of increasing of (*S*)-lysine concentration. The data are normalized by dividing the apparent rate (v_{app}) by the highest measured rate (v_{con}) of activity. Each data point (△, with dashed line) was measured at least in duplicate, and the error bars show standard deviations.

Another branch point enzyme, DAH7PS from *M. tuberculosis* from the pathway responsible for the biosynthesis of aromatic compounds, was found to be inhibited by all three aromatic amino acids, and combining two gave greater inhibition than the additive effect alone, suggesting synergistic inhibition.[38] Thus, the effect of the amino acids of the aspartate family (listed in chapter one, section 1.2) on the activity of *M. tuberculosis* DHDPS was of interest. DAP, threonine, and methionine were tested and found to cause no significant inhibition of *M. tuberculosis* DHDPS activity, either separately **(Table 2.6)** or in combination with (*S*)-lysine or each other **(Table 2.7)**.

Table 2.6: The effects of aspartate family amino acids on the activity of *M. tuberculosis* DHDPS. Each substrate concentration was held at 0.3 mM.

Amino acid	Concentration (mM)	Relative activity[a]
meso-DAP	1	99 %
	10	87 %
	20	57 %
(*S*)-Threonine	1	104 %
	10	86 %
	20	74 %
(*S*)-Methionine	1	100 %
	10	99 %
	20	96 %

Table 2.7: The effects of various combinations of aspartate family amino acids on *M. tuberculosis* DHDPS. Each substrate concentration was held at 0.3 mM.

(*S*)-Lysine	*meso*-DAP	(*S*)-Threonine	(*S*)-Methionine	Relative activity[a]
5 mM	5 mM			92 %
5 mM		5 mM		97 %
5 mM			5 mM	95 %
	5 mM	5 mM		86 %
	5 mM		5 mM	99 %
		5 mM	5 mM	91 %

[a] Activity determined in duplicate/triplicate by the coupled assay and reported relative to that in the absence of inhibitor.

The absence of significant feedback inhibition in *M. tuberculosis* DHDPS indicates that the regulation of the (*S*)-lysine biosynthetic pathway (if it is indeed regulated in *M. tuberculosis*) is maintained *via* an alternate strategy to inhibition of DAH7PS activity in *M. tuberculosis*. In *E. coli*, the DAP pathway has been proposed to be controlled at both the aspartate kinase and DHDPS catalytic step.[55] The expression of *E. coli* DHDPS is regulated by the level of DAP,[56]

thus regulating the (*S*)-lysine biosynthesis at the genetic level; however, whether such an approach is adopted in *M. tuberculosis* is beyond the scope of this study.

2.5 Biophysical analysis of *M. tuberculosis* DHDPS

Inhibition of self-association into quaternary structure has been suggested as an alternative approach to drug design,[57] and was a key focus of this work. Hence it was important to determine the state of association of *M. tuberculosis* DHDPS in solution. All DHDPS enzymes that have had their function confirmed are homotetramers, except the recently discovered dimeric DHDPS from methicillin-resistant *S. aureus* (*MRSA*) (discussed in chapter one, section 1.5.1).[58] For *M. tuberculosis* DHDPS, two complementary techniques, gel-filtration chromatography and analytical ultracentrifugation (AUC), were employed to examine quaternary structure *via* the determination of solution molecular mass.

2.5.1 Comparison of the two biophysical techniques

In gel-filtration liquid-chromatography, proteins migrate based on their hydrodynamic radius, with larger proteins eluting first. Experiments carried out in this research involved the application of a small volume of concentrated protein to a gel-filtration column, with the buffer volume required for elution (V_e) corresponding to a peak observed by UV absorbance. This can be related to molar mass using standards such as BSA (67 kDa), and ovalbumin (43 kDa), but variation in hydrodynamic properties leads to approximately 10 % uncertainty.

In analytical ultracentrifugation (AUC), the effect of centrifugal force on protein is monitored and fitted using rigorous theory in order to determine its molecular mass and sedimentation coefficient. Inaccuracies caused by non-ideal behaviour can be avoided by using low protein concentrations ($\leq$1 mg.mL^{-1}), and by including 100 mM of electrolyte, such as NaCl.[59] Two experimental approaches, sedimentation velocity (SV) and sedimentation equilibrium (SE), allow for the examination of solution behaviour using hydrodynamic and thermodynamic models, respectively.[60]

The lack of dependence on calibration or assumptions about shape allows for more accurate determination of molar mass with AUC in comparison to gel-filtration chromatography.[61] Contrastingly, the lack of dependence on models in gel-filtration chromatography means the

technique provides an excellent basis of comparison to detect possible artefacts introduced to AUC data during model fitting.

2.5.2 Analytical ultracentrifugation

In sedimentation velocity (SV), the application of a large centrifugal force causes the depletion of macromolecules from the meniscus and results in the formation of a concentration boundary that gradually migrates down the length of the cell.[59,62] The rate of sedimentation in SV is dependent on other forces in addition to the centrifugal force, such as diffusion.[62] Therefore, a hydrodynamic model is used to describe the net effect of these forces and from this can be derived the Svedberg equation, which introduces the sedimentation coefficient (s) to describe the relationship between the rate of sedimentation (u) and the centrifugal force ($\omega^2 r$):

$$s = \frac{u}{\omega^2 r} = \frac{M(1-\bar{v}\rho)}{Nf}$$ **Equation 2.5** [59,63]

Here s is also described in terms of solution properties, where M is molecular mass, $\bar{v}$ is the density of the protein (or partial specific volume), ρ is the density of the solvent, f is the frictional coefficient (related to the shape of the protein) and N is Avogadro's number.[59,63] In SV, the rate of sedimentation is used to determine s, which can be converted into molecular mass with the knowledge of the other solution properties.[59] Both $\bar{v}$ and ρ can be calculated but f needs to be extracted from experimental data, and therefore SV experiments are more susceptible than sedimentation equilibrium experiments to inaccuracies from non-ideal behaviour, including non-heterogeneity and/or self-associating systems.[59]

For SV experiments with *M. tuberculosis* DHDPS, protein sample and reference were centrifuged at 40000 rpm and data were collected every 8 minutes without averaging **(Figure 2.15)**. The partial specific volume ($\bar{v}$) of the wild-type enzyme (0.7402 $mL.g^{-1}$), buffer density (1.005 $g.mL^{-1}$) and viscosity (1.021 cP) were determined using the program SEDNTERP.[64] SV data were fitted to a continuous sedimentation coefficient [$c(s)$] distribution model, using the program SEDFIT.[65] The sedimentation coefficient (s^*) is taken as the ordinate of the peak observed in this distribution and was converted into the standardized sedimentation coefficient ($s°_{20,w}$), using SEDNTERP.[64]

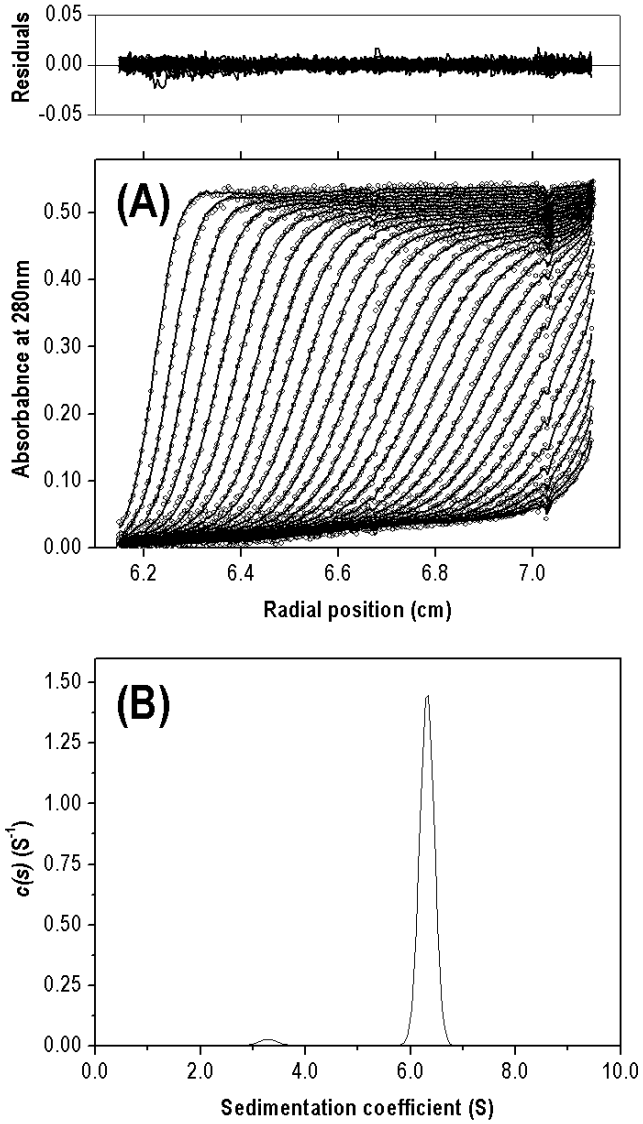


Figure 2.15: **SV analysis of *M. tuberculosis* DHDPS (1.1 mg.mL^{-1}) in 20 mM Tris.HCl, pH 8.0, 150 mM NaCl, at 20 °C. (A) Absorbance at 280 nm plotted as a function of radial position from the axis of rotation (cm). A radial-dependent, time invariant (TI) baseline was subtracted from the data to account for optical imperfections. The raw data (○) are overlaid with the nonlinear least-squares fit (solid line) to a [*c(s)*] model. (B) The *c(s)* distribution was fitted using a resolution of 200 species with $\bar{\nu}$ = 0.7402 mL.g^{-1}, ρ = 1.005 g.mL^{-1}, η = 1.021 cP and f/f_0 = 1.23499. The r.m.s.d. and Z-test for the fit were 0.0038 and 12.17, respectively.**

Data collected at 280 nm and high concentrations of 1.1 mg.mL^{-1} *M. tuberculosis* DHDPS showed one main peak of 6.3 S, with a much a smaller secondary peak of 3.1 S (3 % of signal) **(Figure 2.15B)**. This second peak resulted in a less than ideal fit, as indicated by the

Z-test value of 12.17, but the randomly distributed residuals **(Figure 2.15)** and low root-mean-square difference (r.m.s.d.) indicate the relative goodness of the fit. These values were converted into standardized sedimentation coefficients ($s°_{20,w}$) of 6.5 S and 3.2 S.

Data were also collected at 230 nm and lower concentrations of 0.06 mg.mL^{-1} *M. tuberculosis* DHDPS, and showed a single species in solution and gave a standardized sedimentation coefficient ($s°_{20,w}$) of 6.7 S **(Figure 2.16B)**. The Z-test value of 2.78, the low r.m.s.d. of 0.0039 and random distribution of the residuals indicates the quality of the fit **(Figure 2.16)**. The absence of a second peak at 3.1 S suggests that the peak observed in the higher concentration SV data set does not correspond to a lower order oligomeric species, since that would become more apparent at lower concentrations due to concentration dependent dissociation. It could be that the smaller peak corresponds to a contaminant that absorbs more highly at 280 nm than the 230 nm used for collecting SV data at lower protein concentrations.

Using SEDFIT, the *c(s)* distribution can be converted into a continuous molecular mass [*c(M)*] distribution; however, the accuracy of the distribution depends on the estimate for the frictional ratio (f/f_0) extracted from the experimental data.[59] The *c(M)* distribution gave an apparent molar mass (M_{app}) of 114 and115 kDa for the data sets of 280 and 230 nm, respectively. These values are closer to the predicted molar mass for a tetramer (124.6 kDa) than a dimer (62.3 kDa), and thus indicates that *M. tuberculosis* DHDPS is a tetramer in solution at 1.1 and 0.06 mg.mL^{-1}. Interestingly, at a concentration of 0.05 mg.mL^{-1} *E. coli* DHDPS shows a significant portion of dimer,[57] which suggests that the dimer-dimer interactions for *M. tuberculosis* DHDPS may be stronger than those found in *E. coli* DHDPS.

Since sedimentation coefficients are standardized to conditions corresponding to pure water at 20 °C and extrapolated to zero protein concentration, the $s°_{20,w}$ of 6.7 S obtained from the lower concentration experiment has better accuracy. The $s°_{20,w}$ values for different geometries of tetramer can be predicted using **Equation 2.6**, where $\bar{v}$ is the partial specific volume, M is the predicted molar mass of the tetramer, and F the geometric factor.[66]

$$s_{20,w} = \frac{0.010 M^{2/3}(1-\bar{v}\rho)}{\bar{v}^{1/3}} F \qquad \textbf{Equation 2.6}\ ^{66}$$

Using *F* values determined using hydrodynamic theory, the $s°_{20,w}$ was calculated as 6.5 S for a square-planar tetrameric *M. tuberculosis* DHDPS (*F* = 0.926) and 6.9 S for a tetrahedral tetramer (F = 0.977).[66] The experimentally determined $s°_{20,w}$ of 6.7 S suggests that the tetramer of *M. tuberculosis* DHDPS adopts an intermediate spatial arrangement.

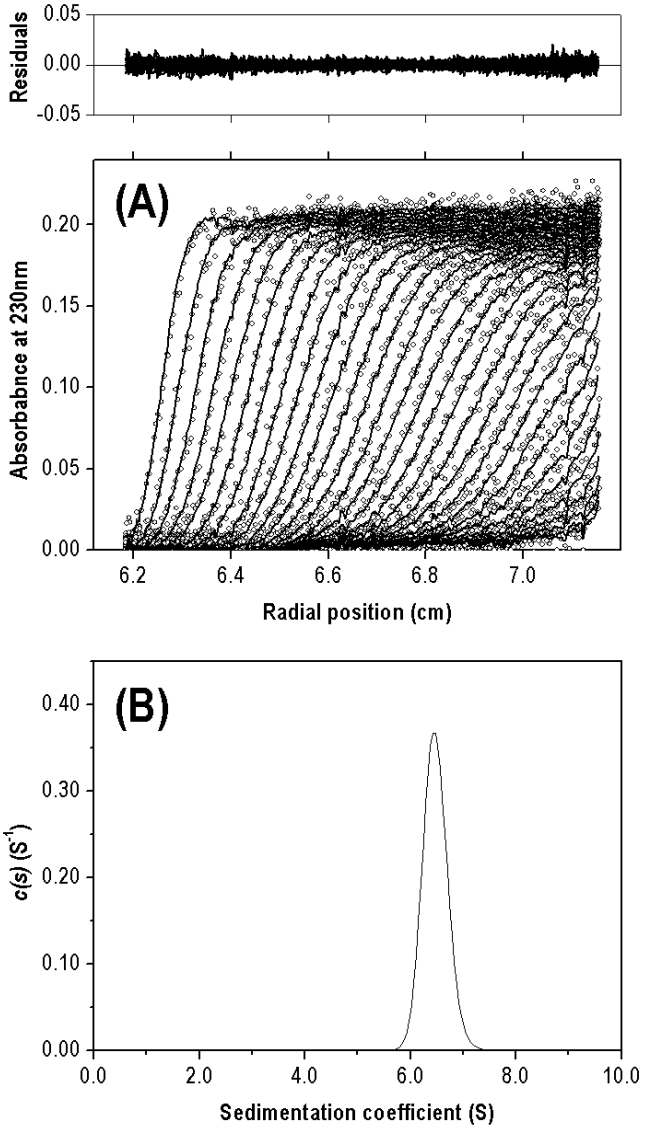


Figure 2.16: **SV analysis of wild-type *M. tuberculosis* DHDPS (0.06 mg.mL^{-1}) in 20 mM Tris.HCl, 150 mM NaCl at 20 °C. (A) The raw data (○) is overlaid with a nonlinear least-squares fit (solid line) to a [*c(s)*] model.[65] (B) The [*c(s)*] distribution is plotted as a function of s^* (in units of Svedberg, S). The fit was obtained using a resolution of 200 species with $\bar{v}$ = 0.7402 mL.g^{-1}, ρ = 1.005 g.mL^{-1}, η = 1.021 cP and f/f_0 = 1.22772 and TI noise was removed. The r.m.s.d. and Z-test for the fit were 0.0039 and 2.78, respectively.**

Wild-type *M. tuberculosis* DHDPS was further analyzed with sedimentation equilibrium (SE) ultracentrifugation, incorporating the results obtained from the SV experiments. In SE, the application of a smaller centrifugal force allows the process of diffusion to significantly oppose the process of sedimentation and given sufficient time, results in the formation of an equilibrium concentration gradient across the length of the cell.[62,67] The thermodynamic model predicts the concentration distributions to adopt an exponential curve, which for the simplest case is governed by molar mass, partial specific volume ($\bar{v}$), and solvent density (ρ).[59,67] Since two of these three quantities can be calculated, or determined experimentally by other means, SE provides the most accurate determination of molar mass. The best results are obtained from global analysis of equilibria formed at multiple rotor speeds, which is particularly important for more complicated cases, such as self-associating systems.[67] In contrast to SV, data analysis requires selecting an association model and thus data analysis is an important and time-consuming step in SE study.[67]

For SE experiments with *M. tuberculosis* DHDPS, samples at initial protein concentrations of 0.1, 0.3, and 0.9 mg.mL^{-1} were centrifuged at 10000 and 16000 rpm and until equilibrium was reached (~24 hours). Radial absorbance scans were performed at 280 nm and the data acquired at both speeds were fitted with the program SEDPHAT,[61] to yield the equivalent molar mass (M_{eq}) assuming a single species and using the $s°_{20,w}$ of 6.7 S obtained in SV experiments. The lowest concentration (0.1 mg.mL^{-1}) data set provided the best fit with a χ^2 value of 0.37, Z-test value of 2.91 and low r.m.s.d. of 0.0031, and gave a molar mass of 116 kDa **(Figure 2.17)**, which is in good agreement with the apparent molar mass of 115 kDa obtained with SV experiments, and thus confirms that *M. tuberculosis* DHDPS adopts a tetrameric quaternary structure.

Similar molar mass of 116 and 108 kDa were determined from the 0.3 and 0.9 mg.mL^{-1} data sets, respectively, which suggests that self-association of the tetramer is not detectable in the concentration range measured. The fits obtained at the higher concentrations had less accuracy (with χ^2 value of 1.2 and 2.9), possibly due to equilibrium not being fully reached; thus, global fitting of all data sets to a single species model gave a M_{eq} of 109 kDa and global reduced χ^2 of 1.6. Additionally, global fitting of the data sets to various self-associating models in SEDPHAT,[61] such as the dimer-tetramer, and monomer-dimer-tetramer model, gave dissociation constants well below the concentration range measured, clearly indicating

that the equilibrium between the oligomeric structures in the self-association of the tetramer was not detectable with this analysis.

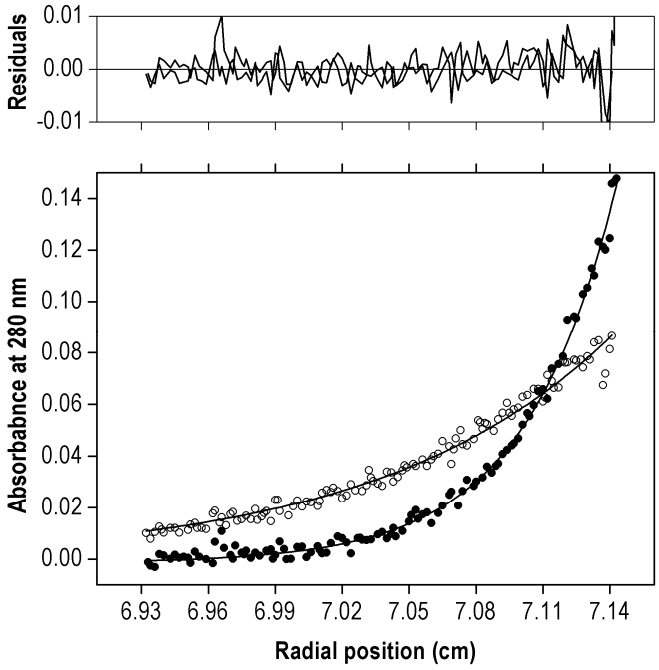


Figure 2.17: **SE analysis of *M. tuberculosis* DHDPS in 20 mM Tris.HCl, 150 mM NaCl, pH 8.0 at 20 °C, with an initial protein concentration of 0.1 mg.mL^{-1}. Absorbance at 280 nm, plotted as a function of radial position from the axis of rotation (cm), after reaching equilibrium at 10000 rpm (○) and 16000 rpm (●). Data are overlaid with the nonlinear least-squares fit (solid line) to a single species model, yielding a molar mass of 116 kDa and χ^2 value of 0.37. The residuals are plotted as a function of radial position from the axis of rotation.**

2.5.3 Gel-filtration liquid-chromatography

In gel-filtration liquid-chromatography, unlike conventional filtration or chromatography methods, proteins are never bound or retained but, instead, separation is based on differing diffusion rates of molecules into the resin pores.[68] Smaller molecules equilibrate into the volumes accessible to them inside the resin beads and consequently travel more slowly down the column, which is reflected in the alternative name of size-exclusion chromatography.[68] This separation based on size allows for the estimating protein molecular masses and thus quaternary structure.[69]

The quaternary structure of *M. tuberculosis* DHDPS was examined using a pre-packed Superdex 200 10/300 column (GE Healthcare) pre-equilibrated with 20 mM Tris.HCl, pH 8.0, and where protein was diluted to ~0.01 mg.mL^{-1} over the course of the chromatography. The calibration plot was generated with four points, using the dimeric and monomeric forms of both BSA and ovalbumin and plotting V_e against the natural log of their molecular masses **(Figure 2.18B)**. Chromatographs were produced by measuring absorbance at 205, 215 and 280 nm as a function of buffer volume, and gave elution peaks with maxima corresponding to V_e **(Figure 2.18A)**. *M. tuberculosis* DHDPS gave a fairly symmetric elution peak indicating one main oligomeric state in solution **(Figure 2.18A)**, which is consistent with the absence of lower order oligomeric species observed in AUC experiments.

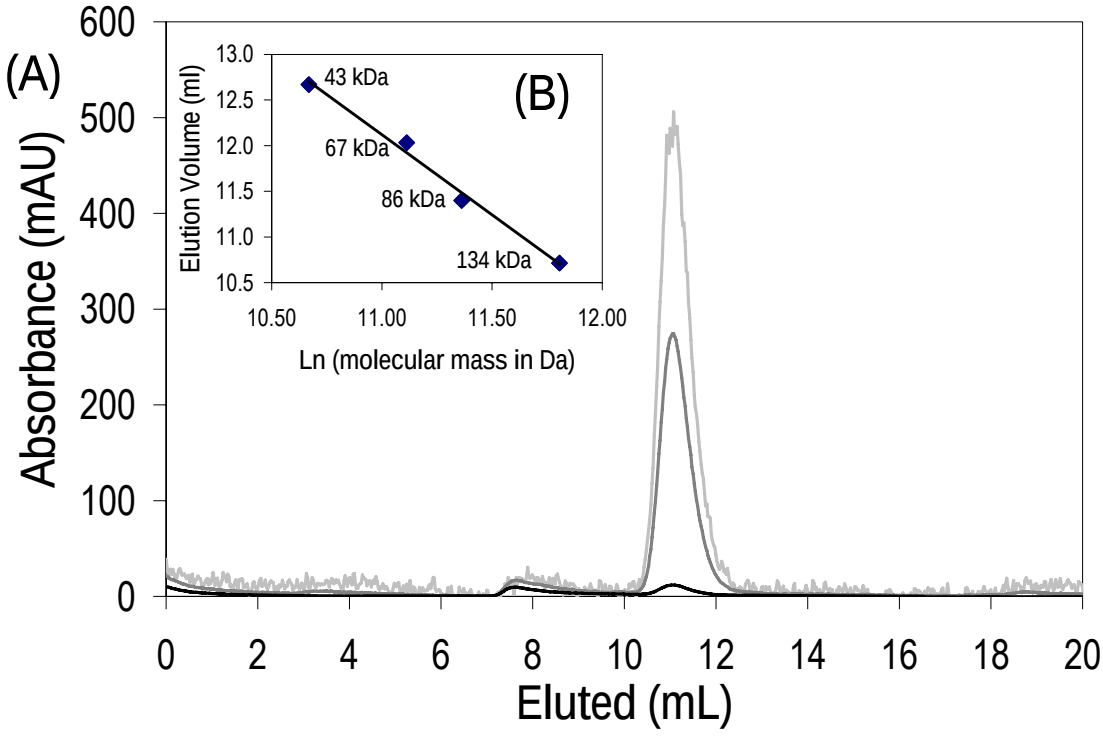


Figure 2.18: (A) Analytical gel-filtration chromatographs of wild-type *M. tuberculosis* DHDPS. The experiment was performed at room temperature in 20 mM Tris.HCl, pH 8.0 and gave an elution peak with a maximum at 10.71 mL as measured by absorbance at 205 nm (light grey). Chromatographs were also recorded at 215 nm (dark grey) and 280 nm (black). (B) BSA and ovalbumin calibration standards (◆) were fitted to a linear equation.

The relationship between V_e and molecular mass (M) was determined from a linear regression of the four point calibration and is shown in the following equation:

$$M = e^{\left[\frac{V_e - b}{a}\right]} \qquad \textbf{Equation 2.7}$$

Here M is related to V_e by the slope, a, and y-intercept, b, of the calibration plot, for this experiment, -1.76 and 31.5 respectively. The molecular mass for *M. tuberculosis* DHDPS was calculated using this equation to be 105 kDa, which is close to the molar mass of 116 kDa determined with sedimentation equilibrium experiments. It represents a 13 % underestimate of the tetrameric molar mass (124.6 kDa), which is only slightly more than the ±10 % uncertainty expected due to assumptions related to the hydrodynamic radius.[70] Thus the analytical gel-filtration experiment is consistent with the AUC analysis and indicates that *M. tuberculosis* DHDPS adopts a tetrameric quaternary structure in solution.

2.6 The crystal structure of *M. tuberculosis* DHDPS

The majority of work to date on DHDPS has investigated *E. coli* DHDPS, providing an excellent basis of comparison for the numerous DHDPS orthologues being determined as part of genomic projects. Our collaborators, Weiss and Kefala, as part of the TB Structural Genomics Consortium, cloned *M. tuberculosis* DHDPS, and determined and analyzed its three-dimensional structure,[1,3] which shows many similarities to *E. coli* DHDPS. These structural coordinates for *M. tuberculosis* DHDPS, at a resolution of 2.28 Å, have been deposited in the PDB databank (entry 1XXX). Material that is largely the work of our collaborators is referenced to our jointly published journal article,[3] whereas work done in the course of this thesis remains unreferenced. Figures illustrating structural details have been prepared using the program PyMOL,[71] unless otherwise stated.

2.6.1 The overall structure

M. tuberculosis DHDPS crystallizes with two tetramers in the asymmetric unit, each comprised of four identical subunits arranged in D_2-symmetry.[3] This correlates well with the sedimentation coefficient ($s°_{20,w}$) determined with AUC, which predicts an intermediate spatial arrangement between a square-planar and tetrahedral tetramer (section 2.5.2). The

monomer is composed of 300 amino acid residues and contains the two domains. The N-terminal $(\beta/\alpha)_8$-barrel domain (residues 1 to 233) and a C-terminal domain (residues 234 to 300) consisting of α-helices,[3] which are fold features that have been observed in all structurally characterized DHDPS enzymes **(Figure 2.19)**.[58,72-78]

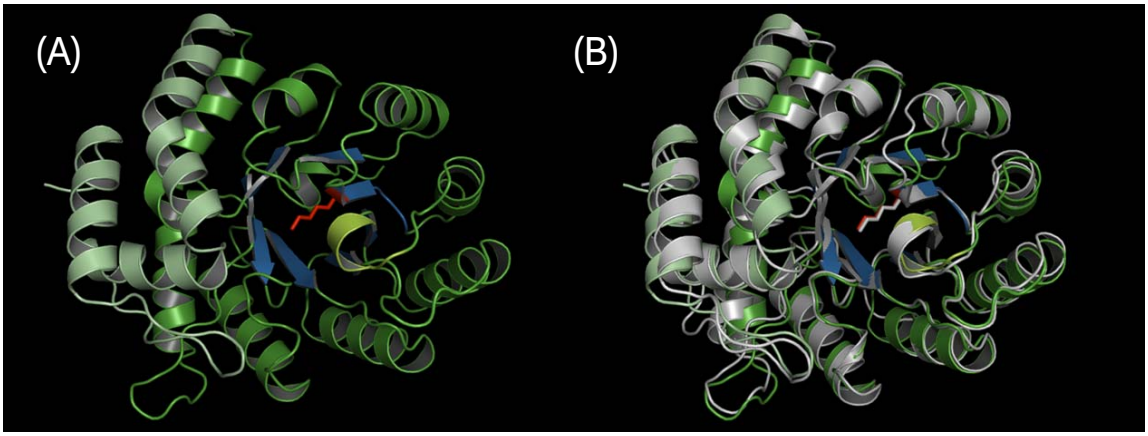


Figure 2.19: The monomeric subunit of (A) *M. tuberculosis* DHDPS has a $(\beta/\alpha)_8$-barrel fold (β-sheets in blue, α-helices in green), which surrounds the active site lysine (red), with an extra α-helix (yellowy green) and a mainly α-helical region (lighter green); (B) the same basic fold is found in *E. coli* DHDPS (overlaid in grey).

The quaternary structure revealed by X-ray crystallography shows *M. tuberculosis* DHDPS existing as a dimer of dimers, with the packing of the "tight-dimer" subunits resembling the bacterial enzyme from *E. coli*, rather than the plant enzyme from *N. sylvestris* **(Figure 2.20)**. The two monomers *a* and *b* **(**and *c* and *d*, **Figure 2.20A)** are tightly bound to form the tight-dimer, with weaker interactions between the *a*/*b* and *c*/*d* tight-dimer units.[3] There are no contacts between subunits *a* and *c* (and *b* and *d*) as there is a large central cavity in the tetramer, like other DHDPS orthologues. Examination of the *M. tuberculosis* DHDPS crystal structure using JavaProtein Dossier suggested that there are more residues involved in the dimer-dimer interface in comparison with the *E. coli* enzyme (to be discussed in more detail in chapter three, section 3.4.1).[79] This seems consistent with the lack of dissociation observed for *M. tuberculosis* DHDPS at high dilutions in both sedimentation velocity ultracentrifugation (0.06 $mg.mL^{-1}$) and gel-filtration (0.01 $mg.mL^{-1}$) experiments (section 2.5.2 and 2.5.3, respectively), in contrast to *E. coli* DHDPS where dissociation was detected at 0.05 $mg.mL^{-1}$.[57]

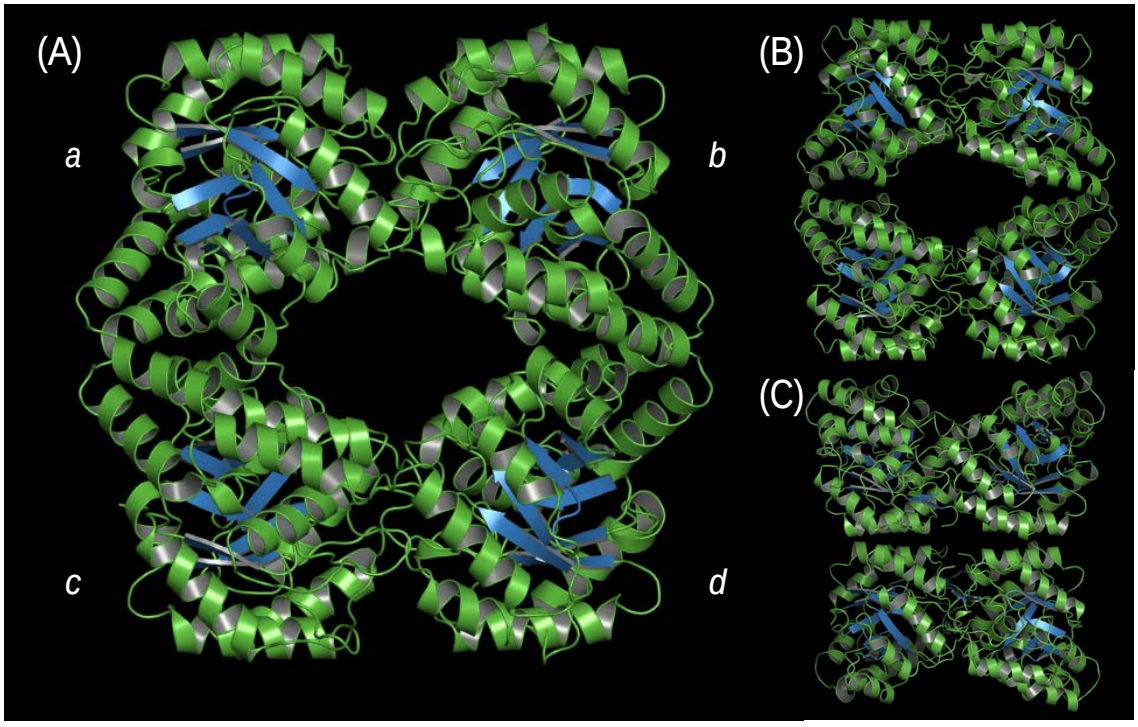


Figure 2.20: The quaternary structure of (A) *M. tuberculosis* DHDPS resembles the dimer of dimers packing of (B) bacterial *E. coli* DHDPS rather than (C) plant *N. sylvestris* DHDPS. The dimer of dimers structure involves subunits *a* and *b* and subunits *c* and *d* associating strongly to form dimers and consequent dimers interacting more weakly to form the tetramer.

The homotetramer of *M. tuberculosis* DHDPS adopts near 222 (D_2) symmetry analogous to that observed in the crystal structures of DHDPS from *Bacillus anthracis*,[77] *C. glutamicum*,[76] *E. coli*,[73] *N. sylvestris*,[74] *T. maritima*,[72] and *Sinorhizobium meliloti*.[78] A least-squares alignment of *M. tuberculosis* DHDPS monomer *a* on these DHDPS, and other structures deposited in the PDB (release date July 2008), was performed using the program LSQMAN,[80] yielded a total of 31 hits, 20 after removing redundancies, with a Z-score above the threshold of 60 **(Table 2.8)**. The pair-wise root-mean-square difference (r.m.s.d.) in α-carbon (Cα) positions between functionally confirmed DHDPS is small, 1.25 Å for *E. coli* DHDPS **(**shown in **Figure 2.19B)**, despite relatively low sequence identity (33 %), confirming that the tertiary structure is conserved.

Of the available structures, *M. tuberculosis* DHDPS had the best structural overlay (r.m.s.d. of 0.60 Å) and highest sequence identity (57 %) with DHDPS from *C. glutamicum* **(Table 2.8)**,

which is not surprising given that *C. glutamicum* is considered a close evolutionary relative to mycobacteria.[53] Interestingly, the relationships between DHDPS and other enzymes of its superfamily, such as NAL,[81,82] KDGA,[83] and recently added YagE DHDPS-like protein,[84] were observed.

Table 2.8: Structural similarities of *M. tuberculosis* DHDPS and other available structures, obtained using LSQMAN from the DEJAVU package.[80] Excludes structures from site-directed mutagenesis studies.

Source	PDB[a] code	Resolution (Å)	Cα's used/total	Sequence identity (%)	R.m.s.d. (Å)	Confirmed DHDPS
C. glutamicum DHDPS	3CPR	2.20	292/301	57	0.60	yes[76]
Bacillus anthracis DHDPS	1XKY	1.94	288/292	40	1.16	yes[77]
Sinorhizobium meliloti DHDPS	2VC6	1.95	282/291	33	1.21	yes[78,85]
M. jannaschii putative DHDPS	2YXG	2.28	281/289	35	1.23	no
Hahella chejuensis putative DHDPS	2RFG	1.50	277/289	35	1.23	no
T. maritima DHDPS	1O5K	1.80	283/294	33	1.24	yes[72]
Aquifex aeolicus putative DHDPS	2EHH	1.90	282/294	33	1.25	no
E. coli DHDPS	1DHP	2.30	282/292	33	1.25	yes[73]
E. coli DHDPS	1YXC	1.90	281/292	33	1.27	yes[55]
S. aureus (*MRSA*) DHDPS	3DAQ	1.45	281/292	32	1.27	yes[58,86]
Clostridium botulinum putative DHDPS	3B18	1.96	280/291	33	1.31	yes[87]
S. aureus (*MRSA*) DHDPS	3DI0	2.38	280/291	32	1.34	yes[75]
E. coli YagE DHDPS-like protein	2V9D	2.15	282/295	25	1.40	no
N. sylvestris[a] DHDPS	-	2.79	273/307	28	1.41	yes
Agrobacterium tumefaciens putative DHDPS	2R8W	1.80	271/292	23	1.52	no
Bacillus clausii putative DHDPS	3E96	1.80	274/295	25	1.55	no
E. coli NAL	1FDY	2.45	279/292	26	1.58	no
Oceanobacillus iheyensis putative DHDPS	3D0C	1.90	273/299	21	1.59	no
Sulfolobus solfataricus KDGA[b]	1W31	1.70	268/293	22	1.62	no
Haemophilus influenzae NAL	1F74	1.60	271/293	23	1.63	no
Rhodopseudomonas palustris putative DHDPS	3DZ1	1.87	265/291	20	1.66	no
Thermus thermophilus putative DHDPS	2PCQ	2.10	259/279	28	1.65	no

[a] Most structural data are freely available from Protein Data Bank (PDB).[88] *Nicotiana sylvestris* DHDPS is not deposited.

[b] *Sulfolobus solfataricus* 2-keto-3-deoxygluconate aldolase (KDGA) is a member of the DHDPS-like/NAL superfamily.

2.6.2 Co-crystallized ligands and metals

The final refined *M. tuberculosis* DHDPS model contains eight DTT molecules covalently bound to cysteine 248 (C248), 8 Mg^{2+}, and 8 Cl^- ions.[3] There is continuous electronic density for DTT extending from the SG of C248, which is found near the dimer-dimer interface.[3] That the activity of *M. tuberculosis* DHDPS with DTT bound was comparable to *E. coli* DHDPS

(section 2.4.4) suggests that this motif is not important for catalysis, though further investigation is needed. Unfortunately, DTT is required during the His_6-tag cleavage reaction; thus, differences in activity observed for *M. tuberculosis* DHDPS unexposed to DTT during purification also reflect the effects of His_6-tag cleavage **(Appendix B)**.

The activity of the enzyme was found to be unaffected by magnesium ions (<20 mM), as determined by the coupled assay. Both the magnesium and chloride ions seen in the structure are likely artefacts of the crystallization conditions and unlikely to have functional significance.[3]

2.6.3 The active site

M. tuberculosis DHDPS contains all active-site residues identified in *E. coli* DHDPS, except asparagine 248 **(Table 2.9, Figure 2.21),** which is replaced by glycine at the equivalent position, a feature previously noted in other Gram-positive bacterial DHDPS isozymes.[21] The mechanism of the DHDPS catalyzed reaction has been studied in detail for the *E. coli* enzyme (discussed in chapter one, section 1.3) by NMR,[21] X-ray crystallography,[55,73] and site-directed mutagenesis,[27,89] and the postulated roles for active-site residues are summarized in **Table 2.9**.

Table 2.9 Identified active-site residues for *E. coli* DHDPS, their equivalent positions in *M. tuberculosis* DHDPS and postulated roles in catalysis from literature.

Active site residues in DHDPS from:		
E. coli	*M. tuberculosis*	Postulated role(s)
Lysine 161	Lysine 171	Binding pyruvate, binding (*S*)-ASA aldehyde
Threonine 45	Threonine 55	Binding pyruvate, binding (*S*)-ASA hydrate
Threonine 44	Threonine 54	Binding pyruvate, proton abstraction (catalytic triad)
Tyrosine 133	Tyrosine 143	Binding pyruvate, proton abstraction (catalytic triad), binding (*S*)-ASA hydrate
Tyrosine 107	Tyrosine 117	Proton abstraction (catalytic triad)
Isoleucine 203	Isoleucine 211	Proton abstraction
Glycine 186	Glycine 194	Binding (*S*)-ASA hydrate, conversion to aldehyde, binding (*S*)-ASA aldehyde
Asparagine 248	not conserved	Binding (*S*)-ASA hydrate
Aspartate 187	Aspartate 195	Binding (*S*)-ASA hydrate/aldehyde
Arginine 138	Arginine 148	Binding (*S*)-ASA hydrate/aldehyde, stabilizing catalytic triad

The active site of DHDPS occurs near the tight-dimer interface, in a pocket created by the association of monomers into a tight-dimer subunit (Figure 1.10, chapter one, section 1.5.3). A key feature of the active site is a catalytic triad involving two tyrosine (Y143, Y117) and a threonine (T54), which together form a proton-relay network, thought to supply and remove

protons at various steps of the reaction catalyzed by DHDPS.[21] Y117 is the only active-site residue not contained within the monomer, but rather reaches in from the adjacent monomer across the tight-dimer interface. All residues, except Y117, fell within the most favoured or additionally allowed regions of the Ramachandran plot for *M. tuberculosis* DHDPS.[3] Similar conformational strain has been observed at the equivalent residue in *E. coli* and other DHDPS enzymes,[21,55,58,72-78] and is consistent with its postulated catalytic role.[90] Both the formation of the active-site pocket and completion of the proton relay provided a rationale for the association of monomers into the tight-dimer subunit in *M. tuberculosis* DHDPS (in a similar manner to *E. coli* DHDPS, as discussed in chapter one, section 1.5.3).

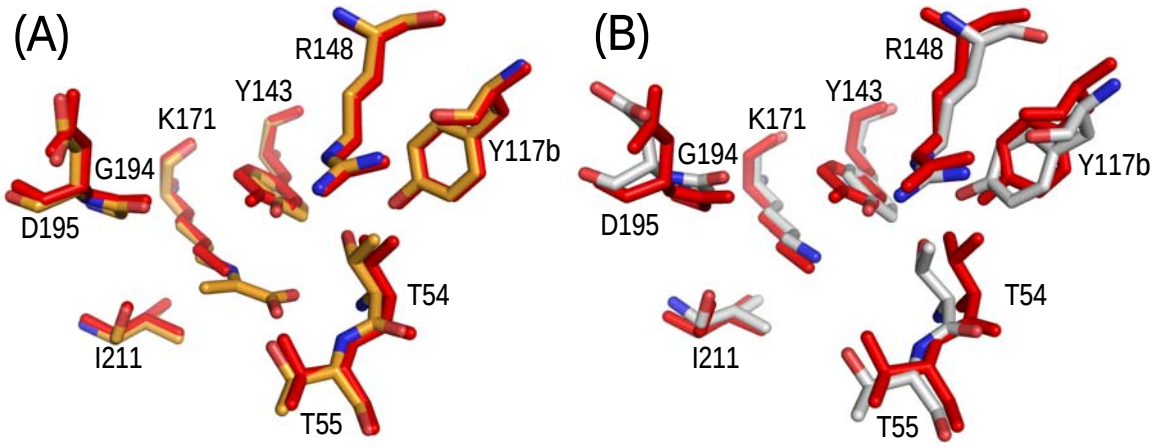


Figure 2.21: Overlays of the key active-site residues of wild-type *M. tuberculosis* DHDPS (red). (A) Comparison of *M. tuberculosis* DHDPS without and with pyruvate bound (in orange) under the same crystallizing conditions shows slightly altered conformation for Y143. (B) Comparison of *M. tuberculosis* DHDPS with the equivalent active-site residues found in *E. coli* DHDPS (in grey, PDB 1XYC) shows several residues adopting slightly altered conformations.

Although the overall architecture of the active site is largely conserved **(Figure 2.21B)**, the proton relay network comprising Y143-OH, T54-OH, and Y117-OH, appears to be disrupted.[3] However, soaking experiments with pyruvate and *M. tuberculosis* DHDPS crystals found a shift of Y143 upon pyruvate binding **(Figure 2.21A)**, and resulted in the distance between Y143-OH and T54-OH shrinking to 3.0 Å **(Table 2.10)**, thus restoring the proton relay network. This shrinking of distances upon pyruvate binding, possibly to facilitate proton transfer, was also observed in crystal structures of *MRSA* DHDPS, suggesting that the catalytic triad is triggered into operation upon pyruvate binding.[75] This has been postulated to

provide a structural basis for the order of substrate binding, as described by the ping-pong reaction mechanism.[75]

Table 2.10 **Comparison of key atomic distances from structures solved for *M. tuberculosis* DHDPS and *E. coli* DHDPS, with and without pyruvate bound (for *E. coli* PDB entries 3DU0 and 1XYC).**

Atomic distances for *M. tuberculosis* DHDPS				Atomic distances for *E. coli* DHDPS			
		no pyruvate	pyruvate[a]			no pyruvate	pyruvate
Y143-OH	K171-Nε[b]	3.3 Å	3.3 Å	Y133-OH	K161-Nε[b]	2.9 Å	3.5 Å
Y143-OH	T54-OH	4.7 Å	3.0 Å	Y133-OH	T44-OH	2.6 Å	2.8 Å
T54-OH	Y117-OH	2.3 Å	2.6 Å	T44-OH	Y107-OH	2.6 Å	2.6 Å
Y117-OH	Y116-OH	6.4 Å	6.5 Å	Y107-OH	Y106-OH	8.5 Å	8.2 Å
Y116-OH	Y90-OH	2.6 Å	2.5 Å	Y116-OH	N80-OH	3.7 Å	3.7 Å

[a] Structure of *M. tuberculosis* DHDPS with pyruvate bound DHDPS has not yet been deposited in the PDB.

[b] The ε-amino group of the key lysine residue that forms a Schiff base with pyruvate.

Another difference in the catalytic triad is that the phenyl ring of Y117 is twisted relative to the equivalent residue in *E. coli* DHDPS **(Figure 2.21B)**. A similar twist has been observed in *MRSA* DHDPS, and was postulated to stabilize the catalytic triad by providing closer aromatic stacking.[75] In *M. tuberculosis* DHDPS, this rotation decreases the distance between Y117 and another tyrosine, Y116 **(Table 2.10)**. The re-orientation of Y117 may result from an additional tyrosine (Y90), in close proximity to Y116 **(Table 2.10)**, which replaces asparagine (N80) in *E. coli* DHDPS **(Figure 2.23)**. Y90 occurs at the tight-dimer interface of *M. tuberculosis* DHDPS and forms aromatic stacking interactions with its counterpart on the opposite subunit. In proteins, aromatic residues forming π interactions commonly adopt either T-shaped or parallel-displaced geometry.[91,92] Y90 together with Y116 and Y117 from both subunits forms a cluster of six aromatic residues. The orientational effects of π stacking have been found to be less distinct in clusters as opposed to pairs of aromatic side chains;[92] thus the additional π orbital interactions of Y90 may be responsible for the altered orientation of both Y116 and Y117 with respect to the equivalent residues in other DHDPS crystal structures.

2.6.4 The (*S*)-lysine binding sites

M. tuberculosis DHDPS was shown to be insensitive to (*S*)-lysine at physiologically relevant concentrations (section 2.4.5). This fits nicely with the observation that most of the (*S*)-lysine

binding site residues are not conserved in *M. tuberculosis* DHDPS, when compared to sequences of DHDPS enzymes that are known to be inhibited by (*S*)-lysine **(Figure 2.22)**. In addition, structural superposition of *M. tuberculosis* and *E. coli* DHDPS **(Figure 2.20)** clearly shows that the (*S*)-lysine can no longer bind in same orientation in the *M. tuberculosis* enzyme. Thus, the further advantage of tight-dimer subunit in *E. coli* DHDPS of providing an allosteric binding site for (*S*)-lysine is not applicable in *M. tuberculosis* DHDPS.

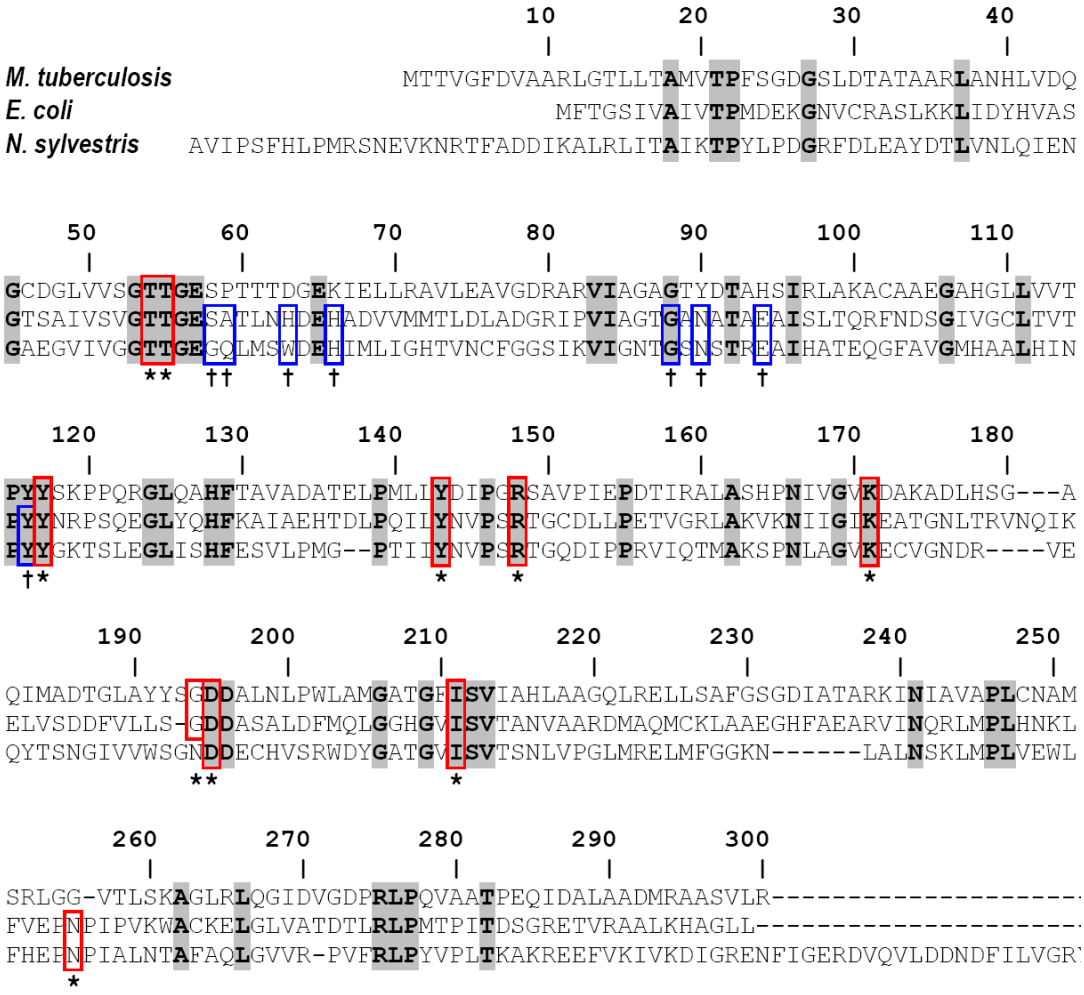


Figure 2.22: **Multiple sequence alignment of the DHDPS sequences from *M. tuberculosis*, *E. coli*, and *N. sylvestris*, using the programs CLUSTALW.[93] The amino-acid residues in grey boxes are conserved in all 3 sequences. The active-site residues are indicated in red boxes (*), while the residues that in *E. coli* and *N. sylvestris* are involved in the (*S*)-lysine binding are indicated by blue boxes (†).**

As noted previously, a tyrosine (*M. tuberculosis*: Y90) replacing asparagine (*E. coli*: N80) seems to extend the aromatic stack of Y116 and Y117 **(Figure 2.20)**. This is particularly interesting as detailed crystallographic study and site-directed mutational analysis have suggested that the (*S*)-lysine binding site in *E. coli* DHDPS is linked to the active site *via* this aromatic stack (*E. coli*: Y106 & Y107, *M. tuberculosis:* Y116 & Y117).[27] In *E. coli* DHDPS, Y106 changes conformation when lysine binds, which alters the orientation of Y107, possibly disrupting its critical role in shuttling protons to and from the active site. Thus the vestigial allosteric cleft of *M. tuberculosis* DHDPS may still provide a binding site for inducing inhibition; although, the site is more shallow than the equivalent site of the *E. coli* DHDPS and therefore precludes (*S*)-lysine from binding **(Figure 2.23)**, as is consistent with the lack of (*S*)-lysine inhibition observed in biochemical characterization (section 2.4.5).

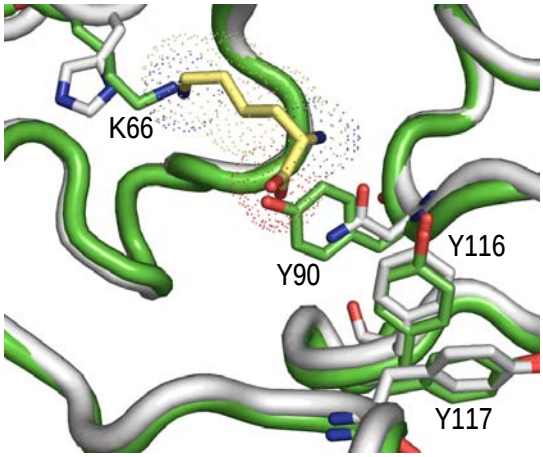


Figure 2.23: Overlays of *M. tuberculosis* DHDPS (in green), with lysine-bound *E. coli* DHDPS (in grey).[55] The residues K66 (*E. coli*: H56) and Y90 (*E. coli*: N80) spatially impede on the position of (*S*)-lysine (in yellow). In *M. tuberculosis* DHDPS, the tyrosine residues, Y90, Y116 and Y117, seem to adopt a parallel-displaced geometry relative to each other, forming an extended aromatic stack.

2.7 Inhibitory studies with *M. tuberculosis* DHDPS

The DAP/(*S*)-lysine biosynthetic pathway has yet to be exploited clinically as a target for antibacterial agents, and several reviews over the past decade have described our increasing knowledge of the enzymes of this pathway as antibiotic targets.[94-98] As discussed in chapter one, section 1.2.1, the DAP biosynthetic pathway has been shown to be essential for *M. tuberculosis*, as has the gene for DHDPS, which catalyzes the pathway's first committed

step,[99] validating the enzyme DHDPS as an antibiotic target.[97] In conjunction with this study, further work has been done by our collaborator Dr Voula Mitsakos to determine the effect of product analogues, initially designed to inhibit *E. coli* DHDPS, on the activity of DHDPS from *M. tuberculosis*,[5] using protein purified during the course of this thesis.

Several compounds **(1a & 1b, Table 2.11)** showed variation in inhibition between *E. coli* and *M. tuberculosis* DHDPS,[5] which highlights the importance of studying orthologues from pressing antibacterial targets, such as *M. tuberculosis*, rather than simply focusing on *E. coli* as a model system. Interestingly, the previous studies with *E. coli* DHDPS indicated that a few of these product analogues **(2a, 3a, 3b, Figure 2.24)** were interacting somewhere other than the active site,[100] and it was postulated that they may be interacting at the (*S*)-lysine binding site.[4] The identical inhibition effect seen in *M. tuberculosis* DHDPS **(2a, 3a, 3b, Table 2.11)**, may suggest binding in the vestigial allosteric cleft (section 6.6.4).

Table 2.11: Inhibitory action of selected product analogues against *M. tuberculosis* DHDPS from the study published by Mitsakos, *et al*.[5]

Compound	Inhibition of enzyme activity[a]	
	E. coli DHDPS	*M. tuberculosis* DHDPS
1a	49 %	43 %
1b	92 %	24 %
2a	76 %	75 %
2b	5 %	0 %
3a	74 %	73 %
3b	85 %	84 %

[a] Enzyme activity was determined in duplicate using the quantitative coupled assay, detailed in section 2.2.3, and were typically within ±3%. Reported as percentage inhibition in the presence of 20 mM compounds 1-3 in Figure 2.24.

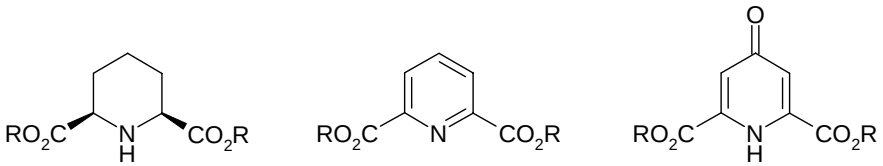


Figure 2.24: Heterocyclic compounds, analogues of the DHDPS product HTPA.[5]

None of these product analogues show themselves to be more potent inhibitors towards *M. tuberculosis* DHDPS as compared to *E. coli* DHDPS **(Table 2.11)**,[4] and the initial study found them to act as weak to moderate inhibitors against *E. coli* DHDPS.[100] This previous study with *E. coli* DHDPS concluded that alternative approaches to design are needed which are not based on substrate or product mimicry to generate sufficiently potent inhibitors to be considered for clinical use.[100] This emphasizes the importance of exploring the emerging paradigm for drug development based on protein-protein interaction disruptors, postulated to be useful antibiotic targets, such as DHDPS, for which no potent active-site inhibitors have yet been found.[101]

2.8 Summary

A comprehensive biophysical, biochemical and structural characterization of wild-type *M. tuberculosis* DHDPS was carried out. The active site of *M. tuberculosis* DHDPS is generally similar to that of *E. coli* DHDPS, with the key residues responsible for substrate binding conserved, suggesting that lead compounds targeted to the *E. coli* active site are likely to be equally effective against the *M. tuberculosis* enzyme. However, collaborative work testing compounds designed to target the active site against *M. tuberculosis* DHDPS did show some differences. Unfortunately, none of these putative inhibitors showed themselves to be potent inhibitors towards *M. tuberculosis* DHDPS, highlighting the importance of exploring the emerging paradigm for drug development based on protein-protein interaction disruptors.

The functional oligomeric unit of the *M. tuberculosis* DHDPS was confirmed to be a tetramer, like the majority of DHDPS orthologues. This was supported by X-ray crystallography as well as several biophysical methods, including analytical ultracentrifugation. The active site requiring a residue, Y117, from the adjacent monomer, provides a rationale for the association of monomers into tight-dimers. The reasons for the association of the tight-dimer subunits into a tetramer are less obvious; and yet dimeric mutants of *E. coli* DHDPS (L197Y and L197D) showed decreased activity and substrate affinity, which indicates quaternary structure is essential for complete catalytic activity.[102] This characterization of *M. tuberculosis* DHDPS provides a basis for comparison in order to understand the effect of disrupted quaternary structure in mutant enzymes, designed to simulate the effect of protein-protein interaction disruptors.

2.9 References

1. Kefala, G. & Weiss, M. S. Cloning, expression, purification, crystallization and preliminary X-ray diffraction analysis of *dap*A (Rv2753c) from *Mycobacterium tuberculosis*. *Acta crystallographica. Section F, Structural biology and crystallization communications* **62**, 1116-9 (2006).

2. Dummler, A., Lawrence, A. M. & de Marco, A. Simplified screening for the detection of soluble fusion constructs expressed in *E. coli* using a modular set of vectors. *Microbial cell factories* **4**, 34 (2005).

3. Kefala, G., Evans, G. L., Griffin, M. D., Devenish, S. R., Pearce, F. G., Perugini, M. A., Gerrard, J. A., Weiss, M. S. & Dobson, R. C. Crystal structure and kinetic study of dihydrodipicolinate synthase from *Mycobacterium tuberculosis*. *The Journal of biological chemistry* **411**, 351-60 (2008).

4. Coulter, C. V., Gerrard, J. A., Kraunsoe, J. A. E. & Pratt, A. J. *Escherichia coli* dihydrodipicolinate synthase and dihydrodipicolinate reductase: kinetic and inhibition studies of two putative herbicide targets. *Pesticide science* **55**, 887-95 (1999).

5. Mitsakos, V., Dobson, R. C., Pearce, F. G., Devenish, S. R., Evans, G. L., Burgess, B. R., Perugini, M. A., Gerrard, J. A. & Hutton, C. A. Inhibiting dihydrodipicolinate synthase across species: towards specificity for pathogens? *Bioorganic & medicinal chemistry letters* **18**, 842-4 (2008).

6. Yugari, Y. & Gilvarg, C. The condensation step in diaminopimelate synthesis. *The Journal of biological chemistry* **240**, 4710-6 (1965).

7. Shedlarski, J. G. & Gilvarg, C. The pyruvate-aspartic semialdehyde condensing enzyme of *Escherichia coli*. *The Journal of biological chemistry* **245**, 1362-73 (1970).

8. Kumpaisal, R., Hashimoto, T. & Yamada, Y. Purification and characterization of dihydrodipicolinate synthase from wheat suspension cultures. *Plant physiology* **85**, 145-51 (1987).

9. Wallsgrove, R. M. & Mazelis, M. Spinach leaf dihydrodipicolinate synthase - partial-purification and characterization. *Phytochemistry* **20**, 2651-5 (1981).

10. Frisch, D. A., Gengenbach, B. G., Tommey, A. M., Sellner, J. M., Somers, D. A. & Myers, D. E. Isolation and characterization of dihydrodipicolinate synthase from maize. *Plant physiology* **96**, 444-52 (1991).

11. Mazelis, M., Whatley, F. R. & Whatley, J. The enzymology of lysine biosynthesis in higher plants. The occurrence, characterization and some regulatory properties of dihydrodipicolinate synthase. *FEBS Letters* **84**, 236-40 (1977).

12. Halling, S. M. & Stahly, D. P. Dihydrodipicolinic acid synthase of *Bacillus licheniformis*. Quaternary structure, kinetics, and stability in the presence of sodium chloride and substrates. *Biochimica et biophysica acta* **452**, 580-96 (1976).

13. Yamakura, F., Ikeda, Y., Kimura, K. & Sasakawa, T. Partial purification and some properties of pyruvate-aspartic semialdehyde condensing enzyme from sporulating *Bacillus subtilis*. *Journal of biochemistry* **76**, 611-21 (1974).

14. Tosaka, O., Ishihara, M., Morinaga, Y. & Takinami, K. Mode of conversion of asparto β-semialdehyde to L-threonine and L-lysine in *Brevibacterium lactofermentum*. *Agricultural and biological chemistry* **43**, 265-70 (1979).

15. Cremer, J., Treptow, C., Eggeling, L. & Sahm, H. Regulation of enzymes of lysine biosynthesis in *Corynebacterium glutamicum*. *Journal of general microbiology* **134**, 3221-9 (1988).

16. Bakhiet, N., Forney, F., Stahly, D. & Daniels, L. Lysine biosynthesis in *Methanobacterium thermoautotrophicum* is by the diaminopimelic acid pathway. *Current microbiology* **10**, 195-8 (1984).

17. Dereppe, C., Bold, G., Ghisalba, O., Ebert, E. & Schar, H. P. Purification and characterization of dihydrodipicolinate synthase from pea. *Plant physiology* **98**, 813-21 (1992).

18. Couper, L., McKendrick, J. E., Robins, David J. & Chrystal, E. J. T. Pyridine and piperidine derivatives as inhibitors of dihydrodipicolinic acid synthase, a key enzyme in the diaminopimelate pathway to L-lysine. *Bioorganic & medicinal chemistry letters* **4**, 2267-72 (1994).

19. Laber, B., Gomis-Ruth, F. X., Romao, M. J. & Huber, R. *Escherichia coli* dihydrodipicolinate synthase. Identification of the active site and crystallization. *The Biochemical journal* **288 (Pt 2)**, 691-5 (1992).

20. Mendelovitz, S. & Aharonowitz, Y. Regulation of cephamycin C synthesis, aspartokinase, dihydrodipicolinic acid synthetase, and homoserine dehydrogenase by aspartic acid family amino acids in *Streptomyces clavuligerus*. *Antimicrobial agents and chemotherapy* **21**, 74-84 (1982).

21. Blickling, S., Renner, C., Laber, B., Pohlenz, H. D., Holak, T. A. & Huber, R. Reaction mechanism of *Escherichia coli* dihydrodipicolinate synthase investigated by X-ray crystallography and NMR spectroscopy. *Biochemistry* **36**, 24-33 (1997).

22. Dixon, M. & Webb, E. C. *Enzymes* (Longman, London, 1979).

23. Scopes, R. K. *Protein purification: principles and practice* (Springer-Verlag, New York, 1987).

24. Cornish-Bowden, A. *Fundamentals of enzyme kinetics* (Butterworths, London ; Boston, 1979).

25. Farkas, W. & Gilvarg, C. The reduction step in diaminopimelic acid biosynthesis. *The Journal of biological chemistry* **240**, 4717-22 (1965).

26. Purich, D. L. *Contemporary enzyme kinetics and mechanism* (Academic Press, New York, 1983).

27. Dobson, R. C., Valegard, K. & Gerrard, J. A. The crystal structure of three site-directed mutants of *Escherichia coli* dihydrodipicolinate synthase: further evidence for a catalytic triad. *Journal of molecular biology* **338**, 329-39 (2004).

28. Dobson, R. C., Griffin, M. D., Roberts, S. J. & Gerrard, J. A. Dihydrodipicolinate synthase (DHDPS) from *Escherichia coli* displays partial mixed inhibition with respect to its first substrate, pyruvate. *Biochimie* **86**, 311-5 (2004).

29. Karsten, W. E. Dihydrodipicolinate synthase from *Escherichia coli*: pH dependent changes in the kinetic mechanism and kinetic mechanism of allosteric inhibition by L-lysine. *Biochemistry* **36**, 1730-9 (1997).

30. Dobson, R. C., Gerrard, J. A. & Pearce, F. G. Dihydrodipicolinate synthase is not inhibited by its substrate, (*S*)-aspartate β-semialdehyde. *The Biochemical journal* **377**, 757-62 (2004).

31. Das, A. Overproduction of proteins in Escherichia coli: Vectors, Hosts, and Strategies in *Guide to protein purification* (ed. Deutscher, M. P.) 93-112 (Academic Press, Sydney, 1990).

32. Cabantous, S., Pedelacq, J. D., Mark, B. L., Naranjo, C., Terwilliger, T. C. & Waldo, G. S. Recent advances in GFP folding reporter and split-GFP solubility reporter technologies. Application to improving the folding and solubility of recalcitrant proteins from *Mycobacterium tuberculosis*. *Journal of structural and functional genomics* **6**, 113-9 (2005).

33. Waugh, D. S. Making the most of affinity tags. *Trends in biotechnology* **23**, 316-20 (2005).

34. Chayen, N. E. Turning protein crystallisation from an art into a science. *Current opinion in structural biology* **14**, 577-83 (2004).

35. Kapust, R. B., Tozser, J., Fox, J. D., Anderson, D. E., Cherry, S., Copeland, T. D. & Waugh, D. S. Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein engineering* **14**, 993-1000 (2001).

36. Fenton, W. A. & Horwich, A. L. GroEL-mediated protein folding. *Protein science: a publication of the protein society* **6**, 743-60 (1997).

37. Goloubinoff, P., Gatenby, A. A. & Lorimer, G. H. GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose bisphosphate carboxylase oligomers in *Escherichia coli*. *Nature* **337**, 44-7 (1989).

38. Webby, C. J., Baker, H. M., Lott, J. S., Baker, E. N. & Parker, E. J. The structure of 3-deoxy-d-arabino-heptulosonate 7-phosphate synthase from *Mycobacterium tuberculosis* reveals a common catalytic scaffold and ancestry for type I and type II enzymes. *Journal of molecular biology* **354**, 927-39 (2005).

39. Novagen. *BL21-CodonPlus® competent cells: instruction manual* (2003).

40. Novagen. *Protein expression, prokaryotic expression: pET system tutorial* (2002).

41. Moffatt, B. A. & Studier, F. W. T7 lysozyme inhibits transcription by T7 RNA polymerase. *Cell* **49**, 221-7 (1987).

42. Studier, F. W. Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. *Journal of molecular biology* **219**, 37-44 (1991).

43. Ericsson, U. B., Hallberg, B. M., Detitta, G. T., Dekker, N. & Nordlund, P. Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Analytical biochemistry* **357**, 289-98 (2006).

44. Deutscher, M. P. Maintaining Protein Stability in *Guide to Protein Purification* (ed. Deutscher, M. P.) 83-9 (Academic Press limited, Sydney, 1990).

45. Cull, M. & McHenry, C. Preparation of extracts from prokaryotes in *Guide to protein purification* (ed. Deutscher, M. P.) 147-53 (Academic Press, Sydney, 1990).

46. QIAGEN. *The QIAexpressionist™ - 5th edition: A handbook for high-level expression and purification of 6xHis-tagged proteins.* (2001).

47. Invitrogen. *AcTEV™ protease: instruction manual* (2005).

48. Pearce, F. G., Sprissler, C. & Gerrard, J. A. Characterization of dihydrodipicolinate reductase from *Thermotoga maritima* reveals evolution of substrate binding kinetics. *Journal of biochemistry* **143**, 617-23 (2008).

49. Eisenthal, R. & Danson, M. J. *Enzyme assays: a practical approach* (IRL Press at Oxford University Press, Oxford [England]; New York, 1992).

50. Fersht, A. *Enzyme structure and mechanism* (W.H. Freeman, New York, 1985).

51. Cornish-Bowden, A. *Fundamentals of enzyme kinetics* (Portland Press, London, 1995).

52. Cornish-Bowden, A. & Wharton, C. W. *Enzyme kinetics* (IRL Press, Oxford [Oxfordshire] ; Washington, D.C., 1988).

53. Pavelka, M. S., Jr. & Jacobs, W. R., Jr. Biosynthesis of diaminopimelate, the precursor of lysine and a component of peptidoglycan, is an essential function of *Mycobacterium smegmatis*. *Journal of bacteriology* **178**, 6496-507 (1996).

54. Fu, L. M. & Fu-Liu, C. S. Is *Mycobacterium tuberculosis* a closer relative to Gram-positive or Gram-negative bacterial pathogens? *Tuberculosis (Edinburgh, Scotland)* **82**, 85-90 (2002).

55. Dobson, R. C., Griffin, M. D., Jameson, G. B. & Gerrard, J. A. The crystal structures of native and (*S*)-lysine-bound dihydrodipicolinate synthase from *Escherichia coli* with improved resolution show new features of biological significance. *Acta crystallographica. Section D, Biological crystallography* **61**, 1116-24 (2005).

56. Acord, J. & Masters, M. Expression from the *Escherichia coli dap*A promoter is regulated by intracellular levels of diaminopimelic acid. *FEMS microbiology letters* **235**, 131-7 (2004).

57. Perugini, M. A., Griffin, M. D., Smith, B. J., Webb, L. E., Davis, A. J., Handman, E. & Gerrard, J. A. Insight into the self-association of key enzymes from pathogenic species. *European biophysics journal* **34**, 469-76 (2005).

58. Burgess, B. R., Dobson, R. C., Bailey, M. F., Atkinson, S. C., Griffin, M. D., Jameson, G. B., Parker, M. W., Gerrard, J. A. & Perugini, M. A. Structure and evolution of a novel dimeric enzyme from a clinically-important bacterial pathogen. *The Journal of biological chemistry* **283**, 27598-603 (2008).

59. Lebowitz, J., Lewis, M. S. & Schuck, P. Modern analytical ultracentrifugation in protein science: a tutorial review. *Protein science: a publication of the protein society* **11**, 2067-79 (2002).

60. Howlett, G. J., Minton, A. P. & Rivas, G. Analytical ultracentrifugation for the study of protein association and assembly. *Current opinion in chemical biology* **10**, 430-6 (2006).

61. Vistica, J., Dam, J., Balbo, A., Yikilmaz, E., Mariuzza, R. A., Rouault, T. A. & Schuck, P. Sedimentation equilibrium analysis of protein interactions with global implicit mass conservation constraints and systematic noise decomposition. *Analytical biochemistry* **326**, 234-56 (2004).

62. Hansen, J. C., Lebowitz, J. & Demeler, B. Analytical ultracentrifugation of complex macromolecular systems. *Biochemistry* **33**, 13155-63 (1994).

63. Ralston, G. *Introduction to analytical ultracentrifugation* (Beckman Instruments, Fullerton, CA, 1993).

64. Laue, T. M., Shah, D. B., Ridgeway, T. M. & Pelletier, S. L. *Analytical ultracentrifugation in biochemistry and protein science* (The Royal Society of Chemistry, Cambridge, 1992).

65. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modelling. *Biophysical journal* **78**, 1606-19 (2000).

66. Teller, D. C., Swanson, E. & de Haen, C. The translational friction coefficient of proteins. *Methods in enzymology* **61**, 103-24 (1979).

67. Balbo, A. & Schuck, P. Analytical ultracentrifugation in the study of protein self-association and heterogeneous protein-protein interactions in *Protein-protein interactions: a molecular cloning manual* (ed. Golemis, E.) 253-78 (Cold Spring Harbor Laboratory Press, New York, 2005).

68. Stellwagen, E. Gel filtration in *Guide to protein purification* (ed. Deutscher, M. P.) 317-28 (Academic Press limited, Sydney, 1990).

69. Winzor, D. J. Analytical exclusion chromatography. *Journal of biochemical and biophysical methods* **56**, 15-52 (2003).

70. Andrews, P. Estimation of the molecular weights of proteins by Sephadex gel-filtration. *The Biochemical journal* **91**, 222-33 (1964).

71. DeLano, W. L. The PyMOL molecular graphics system. (DeLano Scientific, San Carlos, CA, USA, 2006).

72. Pearce, F. G., Perugini, M. A., McKerchar, H. J. & Gerrard, J. A. Dihydrodipicolinate synthase from *Thermotoga maritima*. *The Biochemical journal* **400**, 359-66 (2006).

73. Mirwaldt, C., Korndorfer, I. & Huber, R. The crystal structure of dihydrodipicolinate synthase from *Escherichia coli* at 2.5 Å resolution. *Journal of molecular biology* **246**, 227-39 (1995).

74. Blickling, S., Beisel, H. G., Bozic, D., Knablein, J., Laber, B. & Huber, R. Structure of dihydrodipicolinate synthase of *Nicotiana sylvestris* reveals novel quaternary structure. *Journal of molecular biology* **274**, 608-21 (1997).

75. Girish, T. S., Sharma, E. & Gopal, B. Structural and functional characterization of *Staphylococcus aureus* dihydrodipicolinate synthase. *FEBS letters* **582**, 2923-30 (2008).

76. Rice, E. A., Bannon, G. A., Glenn, K. C., Jeong, S. S., Sturman, E. J. & Rydel, T. J. Characterization and crystal structure of lysine insensitive *Corynebacterium glutamicum* dihydrodipicolinate synthase (cDHDPS) protein. *Archives of biochemistry and biophysics* **480**, 111-21 (2008).

77. Blagova, E., Levdikov, V., Milioti, N., Fogg, M. J., Kalliomaa, A. K., Brannigan, J. A., Wilson, K. S. & Wilkinson, A. J. Crystal structure of dihydrodipicolinate synthase (BA3935) from *Bacillus anthracis* at 1.94 A resolution. *Proteins* **62**, 297-301 (2006).

78. Phenix, C. P., Nienaber, K., Tam, P. H., Delbaere, L. T. & Palmer, D. R. Structural, functional and calorimetric investigation of MosA, a dihydrodipicolinate synthase from *Sinorhizobium*

meliloti l5-30, does not support involvement in rhizopine biosynthesis. *Chembiochem: a European journal of chemical biology* **9**, 1591-602 (2008).

79. Neshich, G., Rocchia, W., Mancini, A. L., Yamagishi, M. E., Kuser, P. R., Fileto, R., Baudet, C., Pinto, I. P., Montagner, A. J., Palandrani, J. F., Krauchenco, J. N., Torres, R. C., Souza, S., Togawa, R. C. & Higa, R. H. JavaProtein Dossier: a novel web-based data visualization tool for comprehensive analysis of protein structure. *Nucleic acids research* **32**, 595-601 (2004).

80. Kleywegt, G. J. & Jones, T. A. Detecting folding motifs and similarities in protein structures. *Methods in enzymology* **277**, 525-45 (1997).

81. Lawrence, M. C., Barbosa, J. A., Smith, B. J., Hall, N. E., Pilling, P. A., Ooi, H. C. & Marcuccio, S. M. Structure and mechanism of a sub-family of enzymes related to N-acetylneuraminate lyase. *Journal of molecular biology* **266**, 381-99 (1997).

82. Babbitt, P. C. & Gerlt, J. A. Understanding enzyme superfamilies. Chemistry as the fundamental determinant in the evolution of new catalytic activities. *The Journal of biological chemistry* **272**, 30591-4 (1997).

83. Theodossis, A., Walden, H., Westwick, E. J., Connaris, H., Lamble, H. J., Hough, D. W., Danson, M. J. & Taylor, G. L. The structural basis for substrate promiscuity in 2-keto-3-deoxygluconate aldolase from the Entner-Doudoroff pathway in *Sulfolobus solfataricus*. *The Journal of biological chemistry* **279**, 43886-92 (2004).

84. Manicka, S., Peleg, Y., Unger, T., Albeck, S., Dym, O., Greenblatt, H. M., Bourenkov, G., Lamzin, V., Krishnaswamy, S. & Sussman, J. L. Crystal structure of YagE, a putative DHDPS-like protein from *Escherichia coli* K12. *Proteins* **71**, 2102-8 (2008).

85. Leduc, Y. A., Phenix, C. P., Puttick, J., Nienaber, K., Palmer, D. R. & Delbaere, L. T. Crystallization, preliminary X-ray diffraction and structure solution of MosA, a dihydrodipicolinate synthase from *Sinorhizobium meliloti* L5-30. *Acta crystallographica. Section F, Structural biology and crystallization communications* **62**, 49-51 (2006).

86. Burgess, B. R., Dobson, R. C., Dogovski, C., Jameson, G. B., Parker, M. W. & Perugini, M. A. Purification, crystallization and preliminary X-ray diffraction studies to near-atomic resolution of dihydrodipicolinate synthase from methicillin-resistant *Staphylococcus aureus*. *Acta crystallographica. Section F, Structural biology and crystallization communications* **64**, 659-61 (2008).

87. Dobson, R. C., Atkinson, S. C., Gorman, M. A., Newman, J. M., Parker, M. W. & Perugini, M. A. The purification, crystallization and preliminary X-ray diffraction analysis of dihydrodipicolinate synthase from *Clostridium botulinum*. *Acta crystallographica. Section F, Structural biology and crystallization communications* **64**, 206-8 (2008).

88. Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. The Protein Data Bank: a computer-based archival file for macromolecular structures. *Journal of molecular biology* **112**, 535-42 (1977).

89. Dobson, R. C., Devenish, S. R., Turner, L. A., Clifford, V. R., Pearce, F. G., Jameson, G. B. & Gerrard, J. A. Role of arginine 138 in the catalysis and regulation of *Escherichia coli* dihydrodipicolinate synthase. *Biochemistry* **44**, 13007-13 (2005).

90. Pearce, F. G., Dobson, R. C., Weber, A., Lane, L. A., McCammon, M. G., Squire, M. A., Perugini, M. A., Jameson, G. B., Robinson, C. V. & Gerrard, J. A. Mutating the tight-dimer

interface of dihydrodipicolinate synthase disrupts the enzyme quaternary structure: toward a monomeric enzyme. *Biochemistry* **47**, 12108-17 (2008).

91. Burley, S. K. & Petsko, G. A. Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science* **229**, 23-8 (1985).

92. McGaughey, G. B., Gagne, M. & Rappe, A. K. π-Stacking interactions. Alive and well in proteins. *The Journal of biological chemistry* **273**, 15458-63 (1998).

93. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research* **22**, 4673-80 (1994).

94. Born, T. L. & Blanchard, J. S. Structure/function studies on enzymes in the diaminopimelate pathway of bacterial cell wall biosynthesis. *Current opinion in chemical biology* **3**, 607-13 (1999).

95. Cox, R. J. The DAP pathway to lysine as a target for antimicrobial agents. *Natural product reports* **13**, 29-43 (1996).

96. Cox, R. J., Sutherland, A. & Vederas, J. C. Bacterial diaminopimelate metabolism as a target for antibiotic design. *Bioorganic & medicinal chemistry letters* **8**, 843-71 (2000).

97. Hutton, C. A., Perugini, M. A. & Gerrard, J. A. Inhibition of lysine biosynthesis: an evolving antibiotic strategy. *Molecular biosystems* **3**, 458-65 (2007).

98. Hutton, C. A., Southwood, T. J. & Turner, J. J. Inhibitors of lysine biosynthesis as antibacterial agents. *Mini reviews in medicinal chemistry* **3**, 115-27 (2003).

99. Sassetti, C. M., Boyd, D. H. & Rubin, E. J. Genes required for mycobacterial growth defined by high density mutagenesis. *Molecular microbiology* **48**, 77-84 (2003).

100. Turner, J. J., Gerrard, J. A. & Hutton, C. A. Heterocyclic inhibitors of dihydrodipicolinate synthase are not competitive. *Bioorganic & medicinal chemistry* **13**, 2133-40 (2005).

101. Gerrard, J. A., Hutton, C. A. & Perugini, M. A. Inhibiting protein-protein interactions as an emerging paradigm for drug discovery. *Mini reviews in medicinal chemistry* **7**, 151-7 (2007).

102. Griffin, M. D., Dobson, R. C., Pearce, F. G., Antonio, L., Whitten, A. E., Liew, C. K., Mackay, J. P., Trewhella, J., Jameson, G. B., Perugini, M. A. & Gerrard, J. A. Evolution of quaternary structure in a homotetrameric enzyme. *Journal of molecular biology* **380**, 691-703 (2008).

Chapter Three

Disrupting quaternary structure with mutagenesis

3.1 Introduction

As described in chapter one, the central goal of this thesis was to investigate the role of quaternary structure in the catalytic activity of DHDPS from *M. tuberculosis*, in order to determine whether tetramer disruption is a viable drug design strategy against this target. DHDPS enzymes typically have strong interactions between pairs of subunits, described as tight-dimers, which associate through weaker interactions to form a tetramer;[1] therefore, the particular focus of this study was the disruption of the weak interface to produce discrete "tight-dimer" subunits.

In this chapter, different point mutations were introduced into the weak interface, building on previous work with the *E. coli* enzyme.[2] Several positions where amino-acid substitution could result in the disruption of the tetrameric structure have been determined based on analysis of the dimer-dimer interface of *M. tuberculosis* DHDPS. It was unclear at the outset of this work whether a single mutation would be sufficient to disrupt the weak interface or whether multiple mutations would be required. Engineering a "dimer" subunit in this way enabled the exploration of the consequence of disrupting quaternary structure on the function of *M. tuberculosis* DHDPS, and thus enabled an assessment of whether interface disruption is a viable strategy for inhibition.

3.2 The importance of quaternary structure

Many enzymes are composed of several subunits in a specific three-dimensional arrangement, or quaternary structure,[3] possibly for reasons of genetic economy,[4] or to increase the size-to-surface ratio of enzymes and lower cellular osmotic pressure.[5] Analysis of observed unfolding pathways has lead to the proposal that unfolded and partially folded monomers may act as chaperones for their partners, guiding proper and rapid assembly, *in vivo*.[6] Oligomerization is suggested to increase protein stability, as a comparison between the stabilization energy (free energy of unfolding) per residue for some monomeric and dimeric proteins shows an exponential decrease as protein size increases.[6] In addition, several thermophilic enzymes

which function under conditions that would denature their mesophilic counterparts have been found to adopt higher oligomeric states than mesophiles.[7-9]

Quaternary structure can play a critical role in enzyme activity.[3,10] Consequently, dissociation into folded subunits, by variation of physical conditions or site-directed mutagenesis, in many cases results in inactivation.[3,10] However, there are some cases, such as carbamoyl-phosphate synthetase, where the monomeric species is more active than the oligomer.[3,11] These changes in activity have mostly been related to catalytic-site formation at the subunit interface or to conformational changes induced by quaternary structure.[3] A small subset of these oligomeric enzymes shows inducible dissociation due to physiologically relevant conditions, such as the binding of a ligand, indicating that quaternary structure also provides the potential for regulatory mechanisms.[3,12]

More recent work has suggested a link between dynamics, quaternary structure and enzyme activity in copper-zinc superoxide dimutase (CuZn-SOD),[13,14] and *E. coli* DHDPS (discussed in chapter one, section 1.5.3).[2] The loss of activity in the disrupted quaternary structure mutants of CuZn-SOD and *E. coli* DHDPS could not be attributed to changes in the active site or tertiary structure; however, increased dynamics were shown by NMR and small angle X-ray scattering, respectively.[2,13,14]

Disrupting wild-type quaternary structure by mutation will be discussed in more detail in the subsequent section 3.3. Several examples, such as the normally heterotetrameric *E. coli* succinyl-CoA synthetase (SCS), show that disrupting quaternary structure does not always adversely affect activity.[15] *E. coli* SCS has been described as a dimer of dimers, like DHDPS, with stable heterodimeric subunits that weakly associate into a tetramer. In contrast to *E. coli* DHDPS, however, a dimeric mutant was found to have comparable activity to the wild-type enzyme, indicating the tetrameric structure was not essential for enzymatic function.[15] It was reasoned that there must be alternative benefits from the association of dimers into tetramers in *E. coli* SCS, such as increased stability and solubility.[15]

Functionality has also been shown to be independent of quaternary structure with monomeric variants of normally homodimeric (or hexameric) insulin,[16,17] and homotetrameric fructose-1,6-bisphosphate,[10] which had activity similar to their wild-type oligomers. In contrast, monomeric variants of triosephosphate isomerase (TIM) were almost inactive.[18-21] The

importance of the homodimeric quaternary structure in the enzyme TIM has been further demonstrated by synthetic peptides which disrupt quaternary structure, thus inhibiting enzyme activity,[22] and there is ongoing research into small molecule dimer disruptors as leads for drug design.[23,24] The observation of inactive, moderately active and fully active monomers, produced by site-directed mutagenesis suggests that quaternary structure plays a variety of roles in nature, including a contribution to enzymatic function in some cases.

3.3 Engineering disrupted quaternary structure

Subunits of multi-subunit proteins have been successfully produced as separate entities using site-directed mutations, both in the exploration of the role of quaternary structure and for industrial and medical purposes. These successes were considered in the design of dimeric variants of *E. coli* DHDPS in the PhD thesis 'Why is dihydrodipicolinate synthase a tetramer?', authored by Griffin,[25] which provided an invaluable reference for this work. There are a limited number of literature examples where the role of quaternary structure has been explored through directed mutation, many of which are discussed in a recent review by Devenish and Gerrard.[26]

The earliest reported designed disruption of quaternary structure involved the homodimeric tyrosyl-tRNA synthetase from *Bacillus stearothermophilus* by Jones *et al*.[27] The residue mutated was among those identified as forming inter-subunit interactions in the native structure, as determined by X-ray crystallography. The mutation of a phenylalanine to aspartate resulted in an enzyme that displayed pH-dependent dissociation into stable, folded monomers due to the electrostatic repulsion between subunits produced by the ionization of the carboxyl group.[27]

However, it is difficult to predict which specific mutations will prevent oligomerization. The approach of introducing charge repulsion encountered difficulties during the design of a dimeric streptavidin, where a histidine to aspartate mutation was predicted to prevent the formation of the natural tetramer.[28] The high hydrophobicity of the dimer-dimer interface caused the mutant streptavidin to form insoluble aggregates rather than discrete, folded dimers. Both successful and unsuccessful attempts to produce characterizable mutants with disrupted quaternary structure are considered in the following sections, from which the principles for designing the mutations in this study were derived.

3.3.1 Loop modification

The protein streptavidin has strong interactions between pairs of subunits, described as stable dimers; however, it normally exists as a tetramer formed from two such stable dimers associating weakly through a small contact area.[28] Investigators were keen to create dimeric streptavidin for industrial purposes, but initial attempts resulted in insoluble aggregates,[28] as discussed in the previous section. Computation and simulations of binding free energy suggested that multiple hydrophobic residues needed to be replaced with polar or charged amino acids to improve solubility, so a more drastic approach was applied of truncating a hydrophobic loop containing 8 residues, which, in combination with the charge repulsion mutation, successfully created soluble, functional, dimeric streptavidin.[28]

In the case of the homodimer TIM, researchers used extensive mutagenesis to create a monomeric variant, by both removing a loop containing 17 of 20 hydrogen bonds formed between the subunits and mutating three hydrophobic residues to increase solubility.[18] The resulting loop-less variant showed no tendency to form dimers or higher aggregates and seemed to adopt a $(\beta/\alpha)_8$-barrel fold similar to wild-type, as reflected by circular dichroism (CD) spectroscopy.[18] The loop-less mutant was 1000-fold less active than wild-type.[18] In contrast, only a 4-fold decrease in activity was observed for monomeric variants of a homodimeric plant enzyme, acetohydroxy acid isomeroreductase, produced by the same loop-deletion approach,[29] suggesting that loop deletion is not always related to dramatic loss of function. However, the extensive mutation involved in loop modification increases the chance of disturbing the active site, as well as disturbing the secondary and tertiary structure of the protein. This provides alternative explanations for the loss of activity observed with the loop-less variant of TIM and undermines the assumption that it is solely disruption of quaternary structure that is responsible for the loss of catalytic activity in this enzyme.

3.3.2 Introducing repulsive forces

Charge repulsion

Closer examination of the mutation choice in the early study with tyrosyl-tRNA synthetase illustrates some important guidelines for rational design.[25,28] The phenyl residue targeted for mutation in this tRNA synthase lies on the symmetry axis of the subunit-subunit interface; therefore, its side chain interacts with the equivalent phenyl group on the other subunit.[27] This

meant that only a single mutation was needed to create the repulsion that separated the subunits, which reflects the principle of seeking minimal mutations to lessen the chance of disturbing the basic structure of the subunit.[25,28] Another guiding principle for engineering of individual subunits from higher-order oligomers is to choose mutations that decrease the hydrophobicity of the interface that becomes exposed upon dissociation into subunits.[25,28] In this early study, the hydrophobic phenyl group was replaced by a hydrophilic aspartate,[27] which presumably increased the solubility of the resultant monomeric variant of tyrosyl-tRNA synthetase.

The strategy of introducing charge repulsion in order to alter quaternary structure was applied to several other systems with varying measures of success. Disruption of oligomerization in the homodimer of CuZn-SOD required a double rather than a single mutation.[13] Charge repulsion was created by the replacement with glutamate of two hydrophobic residues which lay opposite each other on different subunits at the dimerization interface.[13] The mutation successfully produced a completely monomeric CuZn-SOD variant and solution NMR studies linked its lower activity to changes in dynamics as no gross structural changes were observed.[13,14] Contrastingly, double mutations introducing charge repulsion caused significant loss of secondary structure for the homodimeric restriction endonuclease *Eco*R I, whereas a single mutation of leucine to aspartate created a monomer with structure analogous to wild-type.[30] The strategy of introducing electrostatic repulsion by a single mutation was successful in engineering folded monomeric variants from natural dimeric insulin,[31] and dimeric variants from natural tetrameric *E. coli* DHDPS.[2]

Steric repulsion

Monomeric insulin was one of the early success stories of rational design and protein engineering, and has been commercially available since 1996.[16,32] The introduced mutations did not significantly affect the *in vivo* potency of the hormone,[16,17] and yet the monomeric quaternary structure allowed for two to three times faster absorption into the blood stream, providing a medicinal advantage.[31] While the most successful strategy investigated was charge repulsion, the effect of steric hindrance was also examined.[31] The introduction of steric bulk at important residues within the interface was found to disrupt oligomerization to a lesser degree than charge repulsion.[31]

In the construction of monomers from the homodimer malate dehydrogenase (MDH), mutations were based on three factors: (1) removing electrostatic interactions, (2) introducing bad steric contacts, and (3) choosing residues with no direct linkage to the active site.[33] These studies were designed to probe the debated interplay between monomers during catalysis, even though the active sites of MDH are well separated from each other, and distal to the monomer-monomer interface.[33] Analysis of MDH from *E. coli* identified a spatially constricted aspartate interacting with a serine on the opposing subunit.[33] Due to symmetry, mutation of the aspartate to a sterically bulky tyrosine caused steric repulsion at two positions, distal to each other, within the interface.[33] The resulting folded soluble monomer had 14000-fold less activity than wild-type, indicating the importance of the subunit-subunit interface in MDH,[33] and demonstrating steric repulsion as an alternative strategy for preventing oligomerization. This approach was also successfully applied to the creation of dimeric *E. coli* DHDPS by mutation of a central interface leucine residue to tyrosine.[2]

3.3.3 Removal of favourable interactions

Hydrogen bonds, salt bridges and disulfide bonds

Rather than focusing mainly on the introduction of repulsion, other studies have been more interested in the removal of favourable interactions. Monomeric variants have been created by disulfide bond removal through mutation of cysteine residues in several instances.[34,35] In early attempts at producing a monomeric TIM variant a histidine was replaced with an asparagine, and thus the mutation was proposed to disrupt a water-mediated interaction between histidine and an asparagine on the other subunit.[20] Unfortunately, the resulting mixture of dimers and monomers existed in a concentration-dependent equilibrium was not useful for quantitative measurement of the effect of monomerization.[20,21]

The removal of salt bridges in combination with introduced charge repulsion was used to create a monomeric variant for a homodimeric enzyme response regulator, PhoP, from *Bacillus subtilis*.[36] Similarly, in rat prostatic acid phosphatase, the disruption of the homodimer was attributed to the removal of hydrogen bonds and aromatic stacking interactions by two point mutations to charged residues.[37] Mutations were used to invert the polarity of a salt bridge in the homodimer dihydrooratate dehydrogenase from *Lactococcus lactis*, and this resulted in a concentration-dependent monomer-dimer equilibrium.[38]

In all of these examples, the main rationale for the mutation was the importance of the individual residue in inter-subunit interactions. This focus on a single inter-subunit contact is also demonstrated in the exploration of quaternary structure in an extracellular endonuclease from *Serratia marcescens*, where a central interfacial histidine residue, with many inter-subunit hydrogen-bond interactions, was mutated to several different amino acids in order to create a monomeric variant.[39]

Conserved residues and structural homologues

It is difficult to predict the importance of particular interactions based on structural data. The failure of the mutation of a residue central within the interface of TIM to produce an entirely monomeric variant led to the more drastic approach of loop modification, discussed in section 3.3.1. However, the extensive mutation used to create the monomeric TIM variant undermined the assertion that its lack of activity could be directly attributed to the disruption of quaternary structure.[21] This spurred the design of yet another monomeric TIM variant. In contrast to previous attempts, sequence and structural alignments with 45 different TIM enzymes were used to inform design.[21] The analysis revealed several residues on the outer edge of the interface to be highly conserved, and mutating just two of these residues resulted in a monomer with 1000-fold less activity than wild-type, thus providing confirmation of the importance of quaternary structure in the enzymatic function of TIM.[21]

More recently, sequence alignments informed the disruption of oligomerization in a heat shock protein from *M. tuberculosis*.[40] The caveat with this approach is that conserved amino acids at the interface could play an undetermined role in the catalytic mechanism or a specific role in maintenance of the active-site architecture, thus casting doubt as to whether loss of activity can be directly related to quaternary structure.

Another approach to designing altered oligomeric states is comparison with structural homologues of different quaternary structure. Homotetrameric lactate dehydrogenase (LDH) is similar to its dimeric MDH counterparts, as superposition of the structures has shown.[41] A soluble dimeric form of LDH from *Bacillus stearothermophilus* was created by replacing a surface loop with the larger loop found in dimeric MDH.[41]

Less successful was the incorporation of nine residues from a mammalian heterodimeric SCS into the normally heterotetrameric *E. coli* SCS, which resulted in insoluble aggregates.[15] However, a soluble dimeric variant of the *E. coli* enzyme with equivalent activity to the wild-type tetramer was created by mutation of four of the non-conserved residues between bacterial and animal SCS enzymes.[15] This suggests that the dimeric subunits act independently in the catalytic mechanism of tetrameric SCS from *E. coli*.[15]

In the case of the tetrameric biotin-binding proteins, avidin and streptavidin, dimers were designed based on an unrelated but structurally similar family of proteins, fibropellins, specifically those which are found in the extracellular matrix of sea urchin embryos.[42,43] These fibropellins contain several domains including an avidin-like C-terminal domain.[42,43] A stable dimer was obtained for both avidin and streptavidin through mutation of a tryptophan residue to the lysine found at the equivalent position in the fibropellin C-terminal domain, providing a rationale for a seemingly unorthodox substitution.[44,45]

Alanine scanning

Important residues for protein-protein interactions are frequently explored through alanine scanning mutagenesis, a technique which consists of the systematic replacement of non-glycine residues with alanine.[46] The mutations can be viewed as side chain deletions and are usually interpreted as indicating whether or not a particular side chain contributes to association.[46] This method has shown that relatively few residues, referred to as "hot spots", are crucial for association,[47,48] as discussed in section 1.4.2. Obligate oligomers, such as DHDPS, tend to have a larger number of interfacial interactions than more transient complexes that need to associate and disassociate *in vivo*. To compensate for the increased number of residues involved in the interface, studies of quaternary structure in obligate oligomers have used double alanine mutations of residue clusters,[49] alanine mutations guided by molecular modelling,[50] and alanine mutations informed by structural overlay with related species.[51]

3.3.4 Summary of the principles for design

The previous sections have described diverse approaches for rational design of discrete subunits from multi-subunit proteins. Many of these approaches overlap and have often been used in tandem with each other. Although most cases described thus far involved mutation

and disruption of one interface, in the literature there also have been several examples of disrupting multiple interfaces; that is creating monomers from tetramers,[10,44] or hexamers.[31,52] These often involve different strategies for mutation(s) on the different interfaces.

The types of mutations can be broadly grouped into four approaches: removal of interactions, steric repulsion, electrostatic repulsion and loop modification. These are listed in rough order of increasing likelihood to disrupt interface interactions, based on the literature examples reviewed. However, the order also seems to reflect an increased risk of unintentionally disrupting tertiary and secondary structure. Obviously it becomes more difficult to attribute the changes in activity directly to the disruption of quaternary structure given the extensive mutagenesis involved in loop modification, and this is a major disadvantage when using this approach for investigations regarding the role of quaternary structure. Consequently, only strategies based on point mutations were considered for this work.

3.4 Designing mutations for *M. tuberculosis* DHDPS

In designing mutations to disrupt the weak interface of the homotetrameric protein, *M. tuberculosis* DHDPS, this study focused on determining positions where maximum repulsive forces can be introduced, rather than the removal of important interactions. The residues selected for mutation also needed to be remote to the active site, to play no obvious role in the secondary structure and to form inter-subunit interactions through the side chains, as interactions occurring *via* the main chain cannot be simply disrupted by mutation.

3.4.1 Protein-protein interface analysis

In order to design point mutations that might disrupt the association of the dimeric subunits of *M. tuberculosis* DHDPS, the residues involved in the weak interface needed to be determined. Visual examination of wild-type *M. tuberculosis* DHDPS structural data (PDB entry 1XXX) using PYMOL,[53] identified fourteen residues with orientation and proximity (≤4.0 Å) to suggest interaction across the dimer-dimer interface. A total of 32 inter-subunit contacts were determined, as shown in **Figure 3.1,** involving nine hydrophobic and five hydrophilic residues.

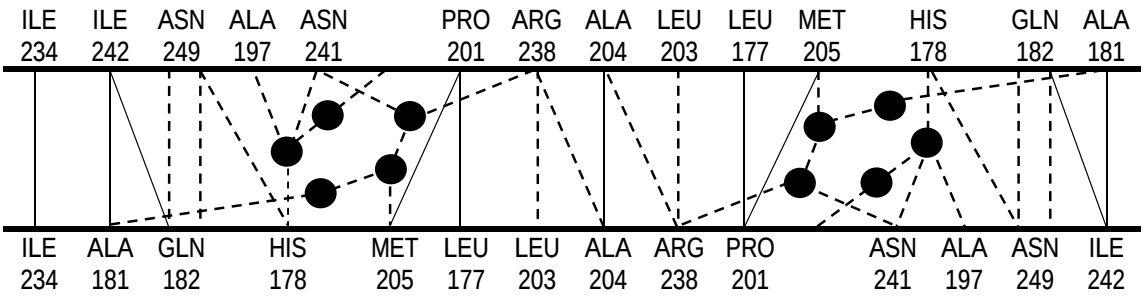


Figure 3.1: **Evaluation of the dimer-dimer interactions of *M. tuberculosis* DHDPS. Hydrogen bonds and charge-charge interactions are indicated by dashed lines with water molecules being represented by dark circles. Hydrophobic and van der Waals contacts are indicated by solid black lines. Contacts were identified by visual inspection, using the program PYMOL,[53] in all sets of weak interfaces in the asymmetric unit.**

The X-ray structure of wild-type *M. tuberculosis* DHDPS contained two tetramers within its asymmetric unit, which provided several sets of weak interface interactions with slight differences, and three additional interactions were identified. Another water near histidine 178 at the dimer-dimer interface allowed for additional water-mediated hydrogen bonds with the backbone carbonyl of aspartate 196 and the backbone NH of asparagine 199. Since this interaction occurs through the main chain and is water-mediated, these residues were not considered as targets for mutation. Of the previously identified interface residues, only leucine 177 (L177) adopted different rotamers and, as a consequence, makes weak hydrophobic contact (3.9 Å) with its counterpart on the opposite subunit in only one of the weak interfaces in the asymmetric unit.

The interface was also examined using stricter criteria with the web-based program JavaProtein Dossier (JDP),[54] which determines inter-subunit contacts in the crystal structure based on maximum distance of 3.2 Å for hydrogen bonds, 6.0 Å for charged (including aromatic) interactions and 3.8 Å for hydrophobic contacts. Several of the interactions first identified were below this threshold, but all previously identified interface residues were predicted to form inter-subunit contacts using JPD **(Figure 3.2**, more details in **Appendix E)**. In addition, five new interface residues were determined: leucine 198, glycine 206, phenylalanine 229, alanine 245 and cysteine 248.

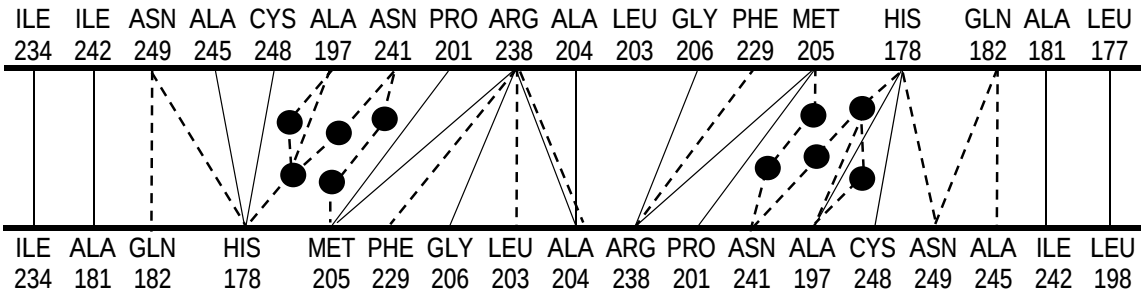


Figure 3.2: Evaluation of the "weak" dimer-dimer interaction of *M. tuberculosis* DHDPS, between chain *a* and *d*, using the program JPD.[54] Hydrogen bonds and charge-charge interactions are indicated by dashed lines with water molecules being represented by dark circles. Hydrophobic and van der Waals contacts are indicated by solid black lines.

Interestingly, asymmetry was apparent in the subunit-subunit interactions **(Figure 3.2)**. For example, the newly determined inter-subunit interaction between L177 and L198 occurred from chain *a* and *d*, but not *vice versa*, presumably due to the different rotamers adopted by L177. Interactions between L177 and its counterpart on the opposite subunit, and with proline 201 were not observed with JPD, due to the shortest atomic distance between the residues falling outside the ≤3.8 Å range.

The dimer-dimer interface was further analysed using the web-based bioinformatics tool, PISA.[55] In this analysis the interface was defined as the protein surface area that becomes inaccessible to solvent when two subunits come into contact,[55] in contrast to previous analyses, which defined the interface as the set of atoms found within the appropriate bonding distance from the other subunit. These two methods for defining the inter-subunit interfaces are widely used in the analysis of macromolecular complexes and have been found to produce roughly equivalent results.[56]

The surface buried by the dimer-dimer interface was 869 (±7) $Å^2$ in *M. tuberculosis* DHDPS **(Figure 3.3B)**, significantly more than the 496 (±2) $Å^2$ buried by the equivalent interface in *E. coli* DHDPS **(Figure 3.3A)**, as determined with PISA.[55] This interface was defined by 24 (±1) residues, more than had been identified by JPD, and involved six non-water-mediated hydrogen bonds of ≤3.2 Å **(**details in **Appendix E)**, which is consistent with the previous analysis using JPD **(Figure 3.2)**. However, only twenty of the interface residues were buried more than 45 % by association of subunits **(**red and orange in **Figure 3.3B)**, which included

eighteen of the interface residues determined by JPD. It was these residues that were considered as candidates for mutation. Similarly, the dimer-dimer interface in *E coli* DHDPS has been previously defined using inter-subunit contacts as nine residues,[57] and only nine of the 15 (±2) interface residues as defined by PISA are buried more than 45 % **(**red and orange in **Figure 3.3A)**.

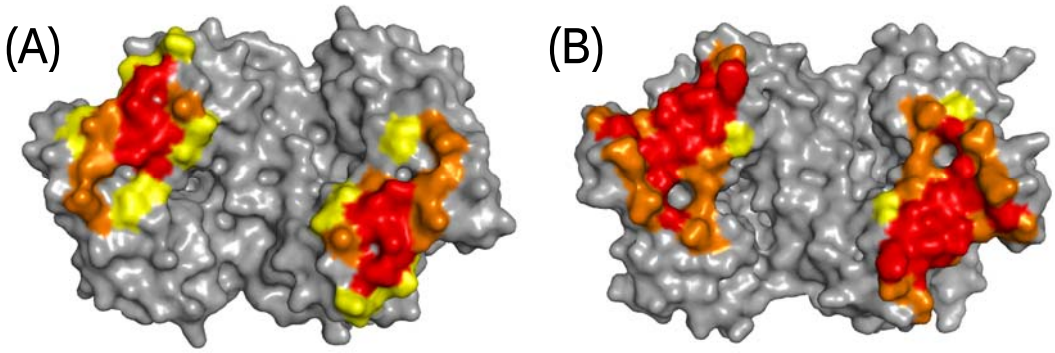


Figure 3.3: Surface view of the dimer-dimer interface from (A) *E. coli* DHDPS and (B) *M. tuberculosis* DHDPS. The residues contributing to the interface are shown in various shades depending on their buried surface area (5 - 44 % yellow, 45 - 84 % orange, ≥85 % red), as defined by PISA.[55]

The greater buried surface area and number of residues involved in the dimer-dimer interface in *M. tuberculosis* DHDPS compared to *E. coli* DHDPS **(Figure 3.3)** suggests that disrupting the weak interface will be more challenging for *M. tuberculosis* DHDPS. It seemed unlikely that a single mutation would disrupt the tetrameric structure of *M. tuberculosis* DHDPS, in contrast to the case with *E. coli* DHDPS. However, it was decided that the first stage of investigation would involve single-point mutations.

3.4.2 Considerations and design of mutations

Conserved residues

Amino-acid sequence alignments for *M. tuberculosis* DHDPS were performed with DHDPS from *E. coli* and *MRSA*. Comparison of the residues involved in the association of dimers into the homotetrameric *E. coli* DHDPS, and the absence of those residues in the naturally dimeric *MRSA* DHDPS, could inform design. Six interface residues of *M. tuberculosis* DHDPS were conserved in the *E. coli* enzyme; although, only leucine (*M. tuberculosis*: L177, *E. coli*: L167)

and arginine (*M. tuberculosis*: R238, *E. coli*: R230) were involved in forming dimer-dimer contacts in both DHDPS enzymes **(Figure 3.4)**.

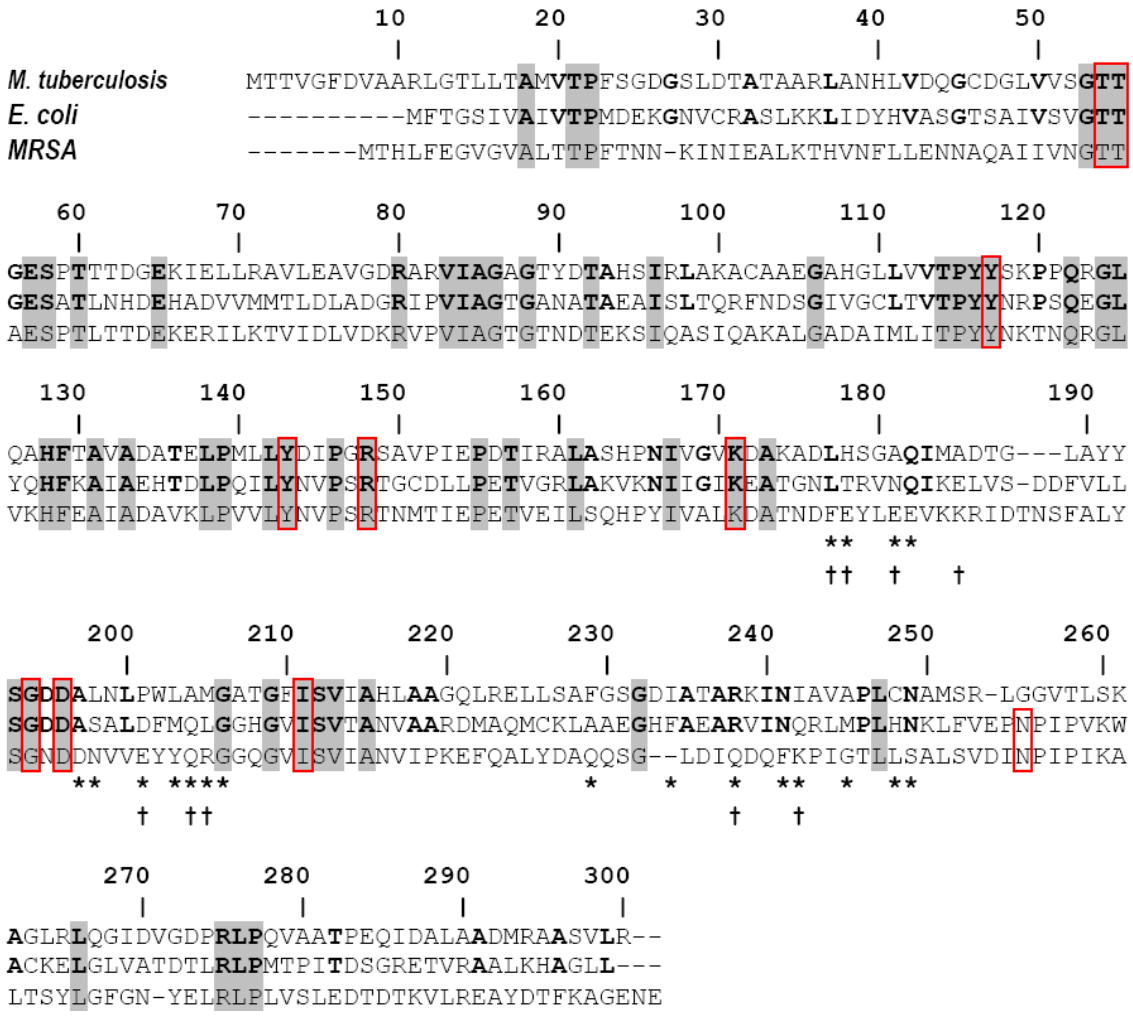


Figure 3.4: **Multiple sequence alignment of the DHDPS sequences from *M. tuberculosis*, *E. coli*, and *MRSA*, using the program CLUSTALW,[58] with numbering corresponding to *M. tuberculosis* DHDPS. The amino-acid residues conserved shown in bold for *E. coli* and *M. tuberculosis* DHDPS, and in grey boxes for all three sequences. The dimer-dimer interface residues with identified inter-subunit contacts are marked with * for *M. tuberculosis* DHDPS and † for *E. coli* DHDPS, while the active-site residues are indicated with red boxes.**

The details of the inter-subunit interactions of these two conserved residues were considered. In *M. tuberculosis* DHDPS, R238 forms inter-subunit contacts with several residues **(**see **Figure 3.2)**, including a hydrophobic interaction with a glycine, which is conserved in *E. coli*

DHDPS (*M. tuberculosis*: G206, *E. coli*: G198); however, the orientation and proximity (>8 Å) precludes the equivalent interaction in *E. coli* DHDPS, and instead R230 forms an interaction with an asparagine that is not conserved in the *M. tuberculosis* enzyme (*E. coli*: N171, *M. tuberculosis*: A181). In both enzymes, the conserved leucine (*M. tuberculosis*: L177, *E. coli*: L167) had been found in crystal structures to make hydrophobic contact with its counterpart on the opposing subunit; however, as mentioned in the previous section, L177 adopts several different rotamers in *M. tuberculosis* DHDPS, disrupting this interaction, whereas, L167 adopts the same rotamer in several crystal structures of *E. coli* DHDPS.

It appears that a distinct set of interactions are involved in the weak interface of DHDPS from *E. coli* and *M. tuberculosis*. The different interactions, even between residues that align based on sequence, highlights the subtle difference in the way the dimers associate in these orthologues. This supports the idea that the association of dimers into a homotetrameric quaternary structure in DHDPS provides a general, rather than a specific, advantage for DHDPS enzymes.[2]

Secondary structure

Only one *M. tuberculosis* DHDPS interface residue, G206, was conserved in *MRSA* DHDPS, the naturally dimeric DHDPS **(Figure 3.4)**. This residue was also conserved in tetrameric *E. coli* DHDPS, although it does not participate in the dimer-dimer interface of this enzyme. This seems to suggest that it may have an additional role, possibly structural, hence it was not considered as a candidate for mutation.

Proline 201 (P201) was also discounted as target for mutation, as proline residues restrict certain peptide torsion angles, and typically play a specific structural role, such as producing slight distortions in the α-helix, as was observed in the crystal structure of *M. tuberculosis* DHDPS. However, it was noted that this residue is not conserved in DHDPS from *E. coli* and *MRSA*, indicating this proline is not a general structural feature in DHDPS enzymes.

Hydrophobicity

The number of hydrophobic residues in the weak interface of *M. tuberculosis* DHDPS is greater than for *E. coli* DHDPS, with eleven hydrophobic interface residues identified using JPD, as opposed to three in *E. coli* DHDPS (Dr Sean Devenish, *pers. comm.*). Insolubility and

aggregation problems seemed likely upon disruption of *M. tuberculosis* DHDPS given the more extensive hydrophobic surface to be exposed. Therefore, the targets for mutation were narrowed to the eleven hydrophobic residues with identified dimer-dimer interactions, by applying the strategy of hydrophobic to hydrophilic mutations in order to maximize solubility of the dimeric variant. This strategy has been used successfully in other studies (section 3.3.2), including the development of a dimeric variant of *E. coli* DHDPS.

The analysis with JPD provided further criteria to narrow the candidates for mutation, as alanine 197 only formed water-mediated hydrogen bonds through its main-chain oxygen. The remaining hydrophobic interface residues and their hydrophobic inter-subunit contacts are listed in **Table 3.1**.

Table 3.1: Hydrophobic contacts between chains *a* and *d* identified by JPD,[54] for the hydrophobic interface residues considered as candidates for mutation.

Residue (from chain *a*)	Distance[a](Å)	Residue (from chain *d*)
Leucine 177	3.7	Leucine 198
Alanine 181	3.8	Isoleucine 242
Leucine 198	3.7	Leucine 177
Proline 201	3.7	Methionine 205
Alanine 204	3.4	Alanine 204
Alanine 204	3.7	Arginine 238
Methionine 205	3.7	Proline 201
Methionine 205	3.4	Arginine 238
Glycine 206	3.7	Arginine 238
Isoleucine 234	3.5	Isoleucine 234
Isoleucine 242	3.8	Alanine 181
Alanine 245	3.6	Histidine 178
Alanine 245	3.8	Glutamine 182

[a] The smallest atomic distance observed between the residues observed in the crystal structure.

Due to concerns with regard to disrupting secondary structure previously discussed, P201 and G206 were not considered as candidates for mutation. Neither L177, nor its hydrophobic contact L198, were considered as a targets for mutation, as a consequence of the apparent flexibility of L177 reflected in the multiple conformers observed in the crystal structure. This narrowed the candidates for mutation to six residues: alanine 181 (A181), alanine 204 (A204), methionine (M205), isoleucine 234 (I234), isoleucine 242 (I242), and alanine 245 (A245).

Of these candidates for mutation, I234 and A204 had the advantage of forming a hydrophobic interaction with their counterpart on the opposite subunit **(Table 3.1)**. Introduction of symmetrically related single mutation will effectively introduce repulsive forces from both faces of the interaction; hence, a single mutation of a hydrophobic residue is doubly powerful for preventing dimer association.

A204, being a relatively small amino acid and more central within the interface than I234, **(Figure 3.5)** was likely to have tighter steric packing, and therefore could provide more scope for applying the strategy of introducing steric contacts to prevent oligomerization (outlined in section 3.3.2). This strategy was successfully applied in the development of a dimeric variant of *E. coli* DHDPS.[2] Interestingly, the leucine residue (*E. coli*: L197, *M. tuberculosis:* M205), which was mutated to produce dimeric variants of *E. coli* DHDPS, occupied a spatially similar position as A204 within the dimer-dimer interface **(**details in **Appendix F)**. Three different mutations were designed for A204, introducing steric repulsion (tyrosine), electrostatic repulsion (aspartate), and both steric and electrostatic repulsion (arginine) at position 204.

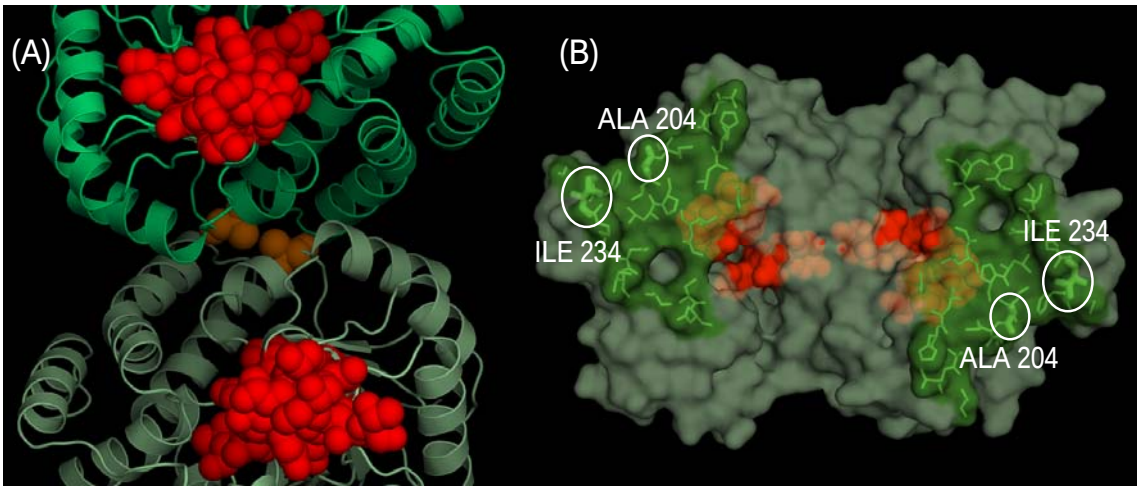


Figure 3.5: **(A) Alanine 204 (in orange) in the weak interface of *M. tuberculosis* DHDPS showing the proximity (3.5 Å) and orientation of this residue between chains *a* (green) and *d* (light green), and distal to the active site (in red). (B) Removing the upper dimeric subunit and looking down onto the weak interface (bright green), chains *c* and *d*, shows the position of A204 and I234 in the weak interface.**

3.5 Site-directed mutagenesis

Site-directed mutagenesis was used to obtain proteins with the specific amino-acid substitutions outlined in the previous section. Mutagenesis was achieved by polymerase chain reaction (PCR) with a double-stranded, circular DNA template, using a commercially available QuikChange® site-directed mutagenesis kit (Stratagene).

3.5.1 Mutagenic primer design

The sequences of the mutagenic primers were designed to introduce mutations to replace alanine at position 204 with tyrosine, aspartate or arginine. These primers were based on the nucleotide sequence of the *dap*A insert carried on the pETM-11 based plasmid designated pMTB02, which was the template for the mutagenic PCR reactions. The sequences of all primers used for mutagenesis are shown in **Table 3.2**.

Table 3.2: **Primer sequences designed for site-directed mutagenesis of *dap*A from *M. tuberculosis* in pMTB02. The codon of the altered amino acid is underlined and the mismatched bases are shown in blue.**

Primer name	Primer sequence
A204Y Forward	5′-CTG CCC TGG CTG TAC ATG GGC GCC AC G-3′
A204Y Reverse	5′-GT GGC GCC CAT GTA CAG CCA GGG CAG-3′
A204R Forward	5′-C CTG CCC TGG CTG CGC ATG GGC GCC ACG-3′
A204R Reverse	5′-CGT GGC GCC CAT GCG CAG CCA GGG CAG G-3′
A204D Forward	5′-CTG CCC TGG CTG GAC ATG GGC GCC AC-3′
A204D Reverse	5′-GT GGC GCC CAT GTC CAG CCA GGG CAG-3′

3.5.2 PCR mutagenesis of plasmid pMTB02

A high purity preparation of pMTB02, isolated from *E. coli* DH5α, was used for the PCR mutagenesis. The primer concentration was kept in large excess of the template, and a control reaction was performed using supplied primers and the control plasmid, pWhitescript, to monitor mutation efficiency.[59] Subsequent to *Dpn* I digestion to remove parental DNA, the presence of mutated plasmid was confirmed using agarose gel electrophoresis. Plasmids were labelled according to the potential mutation carried, e.g. the plasmid from the PCR reaction containing A204D forward and reverse primers was designated pMTBA204D.

3.5.3 Transformation of XL1-Blue with mutated DNA

E. coli XL1-Blue supercompetent cells were transformed with the newly synthesized DNA from the mutagenic PCR reactions, using the heat-shock method (chapter six, section 6.2.12). Transformed cells were selected and cultured on LB agar plates containing kanamycin. From these plates, four potential transformant colonies for each mutation were sub-cultured and stored as glycerol stocks at -80 °C. The mutation efficiency of the PCR conditions was assessed using the control plasmid, pWhitescript,[59] which gave 74 % blue colonies. Thus the proportion of mutated plasmids was similar to the 80 % efficiency expected.

3.5.4 Sequencing of mutated *dap*A genes

The presence of the mutation and the integrity of the full-length *dap*A insert after mutagenesis were determined using bi-directional sequencing performed by Canterbury Sequencing at the University of Canterbury. The procedure was based on the dideoxynucleic acid chain termination method,[60] with DNA samples prepared to a high purity by standard methods. Both strands for each mutation of all four potential transformants were examined and the best sequence reads gave 700 to 900 bp. This allowed a full length consensus sequence to be determined for each mutation, which confirmed the desired mutations **(Figure 3.6)** and the absence of undesired mutations in the rest of the *dap*A insert.

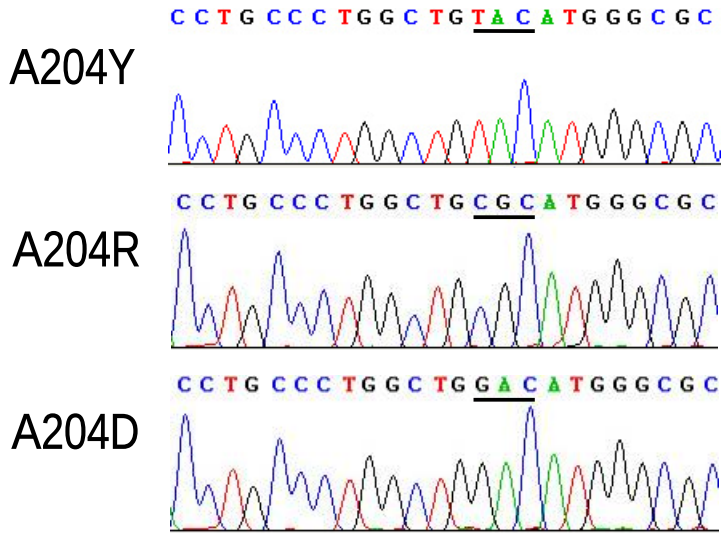


Figure 3.6: **Chromatograms and sequence assignments near the mutation site, with the mutated amino-acid codon underlined. The data shown are the reverse complement of the reverse read, which consistently showed a better signal-to-noise ratio near the mutation site.**

3.6 Analysis of quaternary structure

Small-scale protein preparations were attempted contemporaneously for all three amino-acid substitutions and their quaternary structures were investigated using blue native (BN)-PAGE.

3.6.1 Transformation of BL21 (DE3) pGroESL with mutated DNA

The plasmid DNA that had provided the best read for sequencing was further used to transform *E. coli* BL21 (DE3) pGroESL cells by electroporation. This *E. coli* strain co-expresses the heat chaperonin proteins, GroES and GroEL,[61,62] which had been found to improve the yield of soluble, biologically active wild-type *M. tuberculosis* DHDPS (chapter two, section 2.3.2). Transformed cells containing both pGroESL and mutated DNA were selected using LB agar plates containing chloramphenicol and kanamycin, and glycerol stocks were made for storage at -80 °C.

3.6.2 Small-scale protein purifications

Analogous to the optimized procedure for wild-type *M. tuberculosis* DHDPS, cells were grown overnight at room temperature and then harvested by centrifugation and re-suspended in extraction buffer, prior to cell lysis by ultrasonication, followed with clarification by centrifugation. Three small-scale purifications were performed with cultures over-expressing the interface mutants DHDPS-A204Y, DHDPS-A204R and DHDPS-A204D.

The proteins were purified using affinity and gel-filtration chromatography and the effect of each purification step on the total activity of the fraction was monitored **(Table 3.3)**. Since the interface mutations were likely to decrease enzyme stability, purifications were performed as rapidly as possible, minimizing storage between steps.

The activity of DHDPS-A204Y increased after His_6-tag removal, as observed for the wild-type enzyme (Table 2.2, chapter two, section 2.3.3). The final purified mutant DHDPS enzymes had different specific activities, but all showed an increase from the initial value determined, reflecting an increase in protein purity **(Table 3.3)**. DHDPS-A204D showed the greatest loss of activity over the course of purification. The decrease in total activity and low protein yield observed for DHDPS-A204R and DHDPS-A204D could be related to enzyme

degradation or instability over time; DHDPS-A204Y was significantly more stable than the other mutants.

Table 3.3: Small-scale purifications of *M. tuberculosis* DHDPS variants.

Fraction	Volume (mL)	Protein[a] (mg)	Total activity[a] (units[b])	Specific activity (units[b]/mg)	Relative total activity
DHDPS-A204Y (1 L, 3.93 g of cells)					
Centrifuged (supernatant)	37.5	101	1.2	0.012	
Affinity (His_6-tagged)	16.0	-	1.0	-	82%
Dialysed (out of high imidazole)	15.0		1.9	-	156%
Affinity (Non-His_6-tagged)	16.0	-	2.0	-	169%
Concentrated[c]	0.5	-	-	-	
Gel filtration[c]	4.0	2.1	3.0	1.4	252%
DHDPS-A204R (1 L, 3.73 g of cells)					
Centrifuged (supernatant)	37.5	109	1.3	0.012	
Affinity (His_6-tagged)	9.0	-	1.1	-	81%
Dialysed (out of high imidazole)	10.0	-	0.35	-	27%
Affinity (Non-His_6-tagged)	12.0	-	0.51	-	39%
Concentrated[c]	0.5	-	-	-	
Gel filtration[c]	2.5	0.4	0.28	0.72	21%
DHDPS-A204D (per 1 L, 3.67 g of cells) [d]					
Centrifuged (supernatant)	22.5	34	0.9	0.027	
Affinity (His_6-tagged)	3.0	-	0.8	-	88%
Dialysed (out of high imidazole)	4.0	-	0.26	-	28%
Affinity (Non-His_6-tagged)	7.0	-	0.075	-	8%
Concentrated[c]	0.5	-	-	-	
Gel filtration[c]	1.0	0.25	0.009	0.038	1%

[a] Protein concentration and enzyme activity were determined using the Bradford and coupled assay, respectively.

[b] 1 unit is defined as the consumption of 1 µmol of NADPH per second.

[c] Protein was concentrated prior to gel filtration using an ultrafiltration spin column, described in section 6.4.7.

[d] 2 L preparation; values halved for ease of comparison to other mutations.

3.6.3 Blue native (BN)-PAGE

BN-PAGE separates proteins on a non-denaturing polyacrylamide gel and has been optimized for the analysis of oligomeric structure by the inclusion in the cathode buffer of the anionic

dye, Coomassie blue G-250.[63] Using this electrophoretic technique, the mutated enzymes were compared directly with tetrameric wild-type *M. tuberculosis* DHDPS **(Figure 3.7)**.

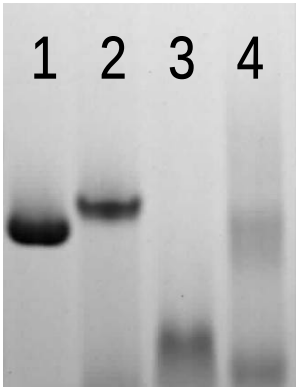


Figure 3.7: BN-PAGE of purified DHDPS-A204Y, DHDPS-A204R and DHDPS-A204D (lanes 2, 3 and 4, respectively), compared with tetrameric wild-type enzyme (lane 1).

DHDPS-A204Y migrated similarly to wild-type *M. tuberculosis* DHDPS, indicating the mutation of alanine to tyrosine to introduce steric bulk was not enough to disrupt the quaternary structure of the enzyme. The slight difference in migration between the wild-type enzyme and DHDPS-A204Y may reflect the latter absorbing less dye than wild-type DHDPS. The non-disrupted quaternary structure of DHDPS-A204Y explains the previously observed activity, yield and stability measurements (in section 3.6.2). This is in contrast to *E. coli* DHDPS, where the L197Y mutation resulted in a completely dimeric species,[2] and suggests some flexibility for rearrangement within the weak interface of DHDPS from *M. tuberculosis*. Consequently, no further analysis was performed with the interface mutant DHDPS-A204Y.

In contrast, the tetrameric structure of *M. tuberculosis* DHDPS was successfully disrupted in both DHDPS-A204R and DHDPS-A204D. In both cases, the bands were less distinct for the interface mutants, possibly reflecting protein degradation. DHDPS-A204R was a single species, presumed to be dimeric, while the two bands apparent for DHDPS-A204D suggested an equilibrium mixture of dimer and tetramer. Thus, the introduction of charge at position 204 caused enough repulsion to disrupt the weak interface, with the arginine (resulting in DHDPS-A204R) the more successful of the two substitutions. The reason for the low activity of DHDPS-A204D in comparison to DHDPS-A204R remains unclear.

3.7 Preliminary analysis of DHDPS-A204R

Before in-depth characterization of DHDPS-A204R was attempted, the integrity of the mutation and the mutant enzyme's secondary structure were investigated using liquid-chromatography mass spectrometry (LCMS) and circular dichroism (CD) spectroscopy, respectively.

3.7.1 Liquid-chromatography mass spectrometry (LCMS)

Mass spectrometry provided a method for confirming the replacement of alanine by arginine in DHDPS-A204R. A freshly prepared sample of *M. tuberculosis* DHDPS-A204R was analyzed using a high-pressure liquid-chromatography (HPLC) system coupled to a mass spectrometer equipped with an electrospray ionization (ESI) probe. Data were collected and processed, using MassLynx 4.0 software, in collaboration with Dr Marie Squire. The molecular mass determined from mass spectrometry was 34248 Da **(Figure 3.8)**, which matched that predicted based on sequence (34247.8 Da), thus further confirming the successful mutation of alanine to arginine.

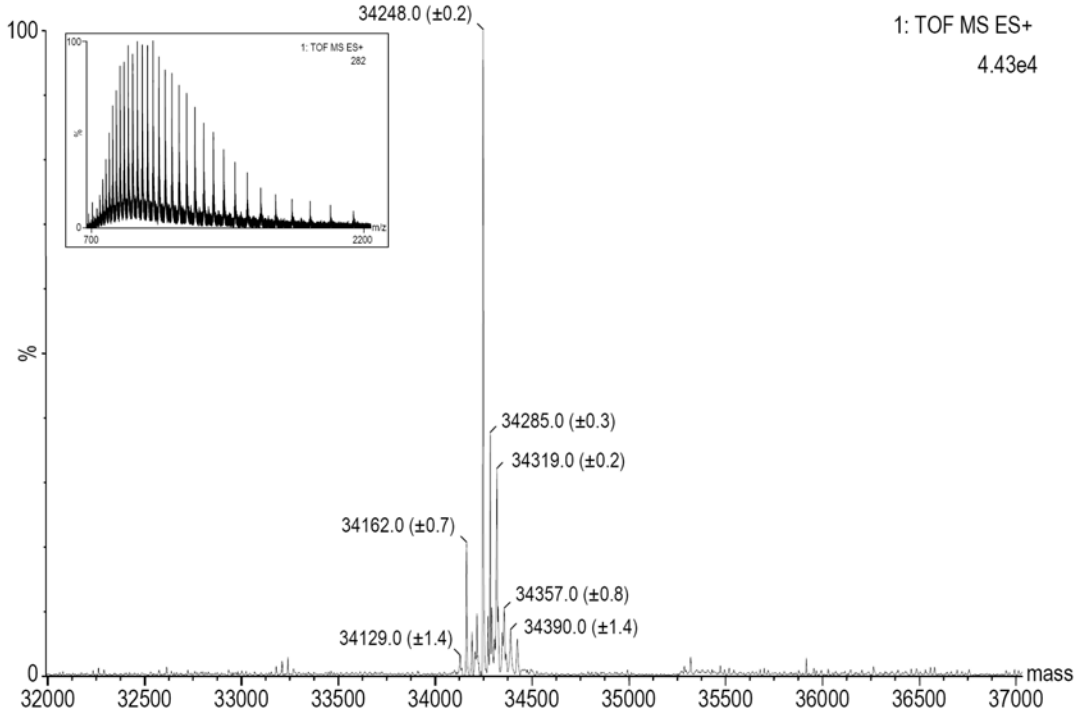


Figure 3.8: LCMS of DHDPS-A204R revealing one main peak which matches the molecular mass predicted from the amino-acid sequence.

3.7.2 Circular dichroism (CD) spectroscopy

Given the apparent instability of DHDPS-A204R, circular dichroism (CD) spectroscopy was employed to confirm that the protein was correctly folded in solution. CD is an important technique for secondary structure determination,[64,65] and monitors the different absorptions of left and right circularly polarized light.[66] Proteins absorb in the far ultraviolet region (180-240 nm) because of their peptide bonds, and thus CD spectra from this region reflect the secondary structure composition of the protein.[67] Each element of the secondary structure gives a characteristic pattern, the most clearly defined being an α-helix with an intense positive peak near 190 nm, and negative peaks at 208 and 222 nm.[64,66] Thus the characteristic double minima of 208 and 222 nm observed in the CD spectra of wild-type and mutant A204R *M. tuberculosis* DHDPS reflects the α-helical components of the structure **(Figure 3.9)**.

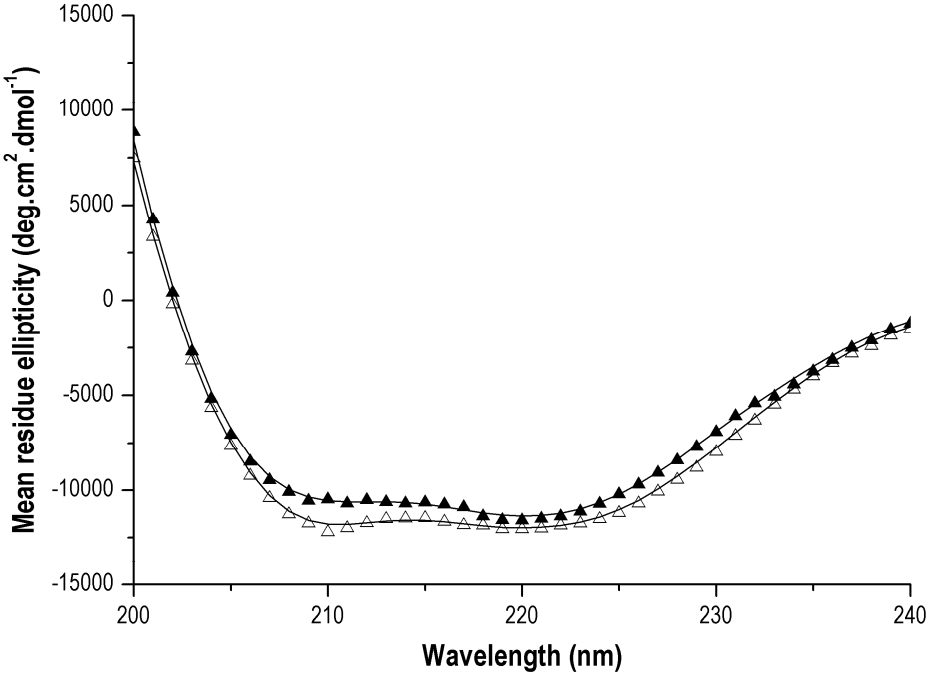


Figure 3.9: The CD spectra of interface mutant DHDPS-A204R (▲) was compared to wild-type *M. tuberculosis* DHDPS (△) from 200 to 240 nm. Data were collected at 0.5 nm intervals at 20 °C, for 150 µg.mL^{-1} protein samples in 20 mM Tris.HCl, pH 8.0. The smoothed data (open symbols) in 1 nm intervals is presented, overlaid with the best fit (solid line), generated by CONTINLL algorithm,[68,69] with the reference set SP29.[70]

The broad minimum between 208 and 222 nm observed for both wild-type and A204R *M. tuberculosis* DHDPS is consistent with the $(\beta/\alpha)_8$-barrel structure of DHDPS, and previous CD spectra observed for *E. coli* DHDPS.[71,72] However, there are differences between the CD spectra of wild-type and mutant DHDPS. To quantify these differences, the data were analyzed using three different algorithms, CONTINLL, SELCON, CDSSTR, in combination with various reference data sets from the CDPro software package.[65,70] The best overall fit using reference set SP29 is summarized in **Table 3.4**. The analysis suggests that the DHDPS-A204R has a slight increase in β-structure. However, these values are not statistically significant and are likely to be artefacts reflecting the limitations of analysis, which has been shown to have variable accuracy for β-structure determination.[66] The majority of the protein was clearly folded in solution in an analogous manner to wild-type. Thus, DHDPS-A204R was selected for structure determination by X-ray crystallography and for full biophysical and biochemical characterization.

Table 3.4: Results from the fits for CD spectra achieved by CONTINLL SELCON, CDSSTR, algorithm,[68,69] with the reference set SP29.[70] The average value of these algorithms is reported, with standard deviation reflecting the difference between these fits.

M. tuberculosis DHDPS	α-Helix	β-Strand	β-Turn	Unordered	R.m.s.d.
Wild-type	38 (±1) %	15.5 (±0.2) %	20.7 (±0.4) %	25 (±1) %	0.04 (±0.02)
Mutant A204R	36 (±2) %	18 (±1) %	21 (±1) %	24 (±2) %	0.052 (±0.001)

3.8 Conclusions

Surprisingly, it appeared that the quaternary structure of *M. tuberculosis* DHDPS could be completely disrupted by a single point mutation. BN-PAGE showed a complete absence of any tetrameric species for the interface mutant DHDPS-A204R. The interface mutations leading to DHDPS-A204Y and DHDPS-A204D were less successful. DHDPS-A204Y was a tetramer, whereas DHDPS-A204D showed disrupted quaternary structure with BN-PAGE, but existed as an equilibrating mixture of two different populations. Thus further experimental work was focused on DHDPS-A204R. Analysis by LCMS confirmed the alanine to arginine mutation and CD spectroscopy showed DHDPS-A204R to be folded in solution with secondary structure analogous to the wild-type *M. tuberculosis* DHDPS. Hence DHDPS-A204R was an excellent candidate for complete biochemical and biophysical characterization in order to investigate the role of tetrameric quaternary structure in the

catalytic activity of *M. tuberculosis* DHDPS. In addition, the fully disrupted quaternary structure of DHDPS-A204R due to a single mutation supports the feasibility of disrupting the homotetramer complex of *M. tuberculosis* DHDPS with a small molecule inhibitor.

3.9 References

1. Blickling, S. & Knablein, J. Feedback inhibition of dihydrodipicolinate synthase enzymes by L-lysine. *Biological chemistry* **378**, 207-10 (1997).

2. Griffin, M. D., Dobson, R. C., Pearce, F. G., Antonio, L., Whitten, A. E., Liew, C. K., Mackay, J. P., Trewhella, J., Jameson, G. B., Perugini, M. A. & Gerrard, J. A. Evolution of quaternary structure in a homotetrameric enzyme. *Journal of molecular biology* **380**, 691-703 (2008).

3. Traut, T. W. Dissociation of enzyme oligomers: a mechanism for allosteric regulation. *Critical reviews in biochemistry and molecular biology* **29**, 125-63 (1994).

4. Goodsell, D. S. & Olson, A. J. Structure symmetry and protein function. *Annual review of biophysics and biomolecular structure* **29**, 105-53 (2000).

5. Perham, R. N. Self-assembly of biological macromolecules. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **272**, 123-36 (1975).

6. Mei, G., Di Venere, A., Rosato, N. & Finazzi-Agro, A. The importance of being dimeric. *The FEBS journal* **272**, 16-27 (2005).

7. Bjork, A., Mantzilas, D., Sirevag, R. & Eijsink, V. G. Electrostatic interactions across the dimer-dimer interface contribute to the pH-dependent stability of a tetrameric malate dehydrogenase. *FEBS Letters* **553**, 423-6 (2003).

8. Clantin, B., Tricot, C., Lonhienne, T., Stalon, V. & Villeret, V. Probing the role of oligomerization in the high thermal stability of *Pyrococcus furiosus* ornithine carbamoyltransferase by site-specific mutants. *European journal of biochemistry* **268**, 3937-42 (2001).

9. Thoma, R., Hennig, M., Sterner, R. & Kirschner, K. Structure and function of mutationally generated monomers of dimeric phosphoribosylanthranilate isomerase from *Thermotoga maritima*. *Structure* **8**, 265-76 (2000).

10. Beernink, P. T. & Tolan, D. R. Disruption of the aldolase A tetramer into catalytically active monomers. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 5374-9 (1996).

11. Lusty, C. J. Catalytically active monomer and dimer forms of rat liver carbamoyl-phosphate synthetase. *Biochemistry* **20**, 3665-74 (1981).

12. Klotz, I. M., Langerman, N. R. & Darnall, D. W. Quaternary structure of proteins. *Annual review of biochemistry* **39**, 25-62 (1970).

13. Bertini, I., Piccioli, M., Viezzoli, M. S., Chiu, C. Y. & Mullenbach, G. T. A spectroscopic characterization of a monomeric analog of copper, zinc superoxide dismutase. *European biophysics journal* **23**, 167-76 (1994).

14. Banci, L., Benedetto, M., Bertini, I., Del Conte, R., Piccioli, M. & Viezzoli, M. S. Solution structure of reduced monomeric Q133M2 copper, zinc superoxide dismutase (SOD). Why is SOD a dimeric enzyme? *Biochemistry* **37**, 11780-91 (1998).

15. Bailey, D. L., Fraser, M. E., Bridger, W. A., James, M. N. & Wolodko, W. T. A dimeric form of *Escherichia coli* succinyl-CoA synthetase produced by site-directed mutagenesis. *Journal of molecular biology* **285**, 1655-66 (1999).

16. DeFelippis, M. R., Chance, R. E. & Frank, B. H. Insulin self-association and the relationship to pharmacokinetics and pharmacodynamics. *Critical reviews in therapeutic drug carrier system* **18**, 201-64 (2001).

17. Hoffman, A. & Ziv, E. Pharmacokinetic considerations of new insulin formulations and routes of administration. *Clinical pharmacokinetics* **33**, 285-301 (1997).

18. Borchert, T. V., Abagyan, R., Jaenicke, R. & Wierenga, R. K. Design, creation, and characterization of a stable, monomeric triosephosphate isomerase. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 1515-8 (1994).

19. Borchert, T. V., Abagyan, R., Kishan, K. V., Zeelen, J. P. & Wierenga, R. K. The crystal structure of an engineered monomeric triosephosphate isomerase, monoTIM: the correct modelling of an eight-residue loop. *Structure* **1**, 205-13 (1993).

20. Borchert, T. V., Pratt, K., Zeelen, J. P., Callens, M., Noble, M. E., Opperdoes, F. R., Michels, P. A. & Wierenga, R. K. Overexpression of trypanosomal triosephosphate isomerase in *Escherichia coli* and characterisation of a dimer-interface mutant. *European journal of biochemistry/FEBS* **211**, 703-10 (1993).

21. Schliebs, W., Thanki, N., Jaenicke, R. & Wierenga, R. K. A double mutation at the tip of the dimer interface loop of triosephosphate isomerase generates active monomers with reduced stability. *Biochemistry* **36**, 9655-62 (1997).

22. Singh, S. K., Maithal, K., Balaram, H. & Balaram, P. Synthetic peptides as inactivators of multimeric enzymes: inhibition of *Plasmodium falciparum* triosephosphate isomerase by interface peptides. *FEBS Letters* **501**, 19-23 (2001).

23. Olivares-Illana, V., Rodriguez-Romero, A., Becker, I., Berzunza, M., Garcia, J., Perez-Montfort, R., Cabrera, N., Lopez-Calahorra, F., de Gomez-Puyou, M. T. & Gomez-Puyou, A. Perturbation of the dimer interface of triosephosphate isomerase and its effect on *Trypanosoma cruzi*. *PLoS neglected tropical diseases* **1**, e1 (2007).

24. Olivares-Illana, V., Perez-Montfort, R., Lopez-Calahorra, F., Costas, M., Rodriguez-Romero, A., Tuena de Gomez-Puyou, M. & Gomez Puyou, A. Structural differences in triosephosphate isomerase from different species and discovery of a multitrypanosomatid inhibitor. *Biochemistry* **45**, 2556-60 (2006).

25. Griffin, M. D. W. Why is dihydrodipicolinate synthase a tetramer? *D. Phil Thesis*, Canterbury. (2005).

26. Devenish, S. R. & Gerrard, J. A. The role of quaternary structure in $(\beta/\alpha)_8$-barrel proteins: evolutionary happenstance or a higher level of structure-function relationships? *Organic and biomolecular chemistry* **7**, 833-9 (2009).

27. Jones, D. H., McMillan, A. J., Fersht, A. R. & Winter, G. Reversible dissociation of dimeric tyrosyl-tRNA synthetase by mutagenesis at the subunit interface. *Biochemistry* **24**, 5852-7 (1985).

28. Sano, T., Vajda, S., Smith, C. L. & Cantor, C. R. Engineering subunit association of multisubunit proteins: a dimeric streptavidin. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 6153-8 (1997).

29. Wessel, P. M., Biou, V., Douce, R. & Dumas, R. A loop deletion in the plant acetohydroxy acid isomeroreductase homodimer generates an active monomer with reduced stability and altered magnesium affinity. *Biochemistry* **37**, 12753-60 (1998).

30. Fritsche, P. & Alves, J. A monomeric mutant of restriction endonuclease *Eco*RI nicks DNA without sequence specificity. *Biological chemistry* **385**, 975-85 (2004).

31. Brange, J., Ribel, U., Hansen, J. F., Dodson, G., Hansen, M. T., Havelund, S., Melberg, S. G., Norris, F., Norris, K., Snel, L., Sørensen, A. R. & Voigt, H. O. Monomeric insulins obtained by protein engineering and their medical implications. *Nature* **333**, 679-82 (1988).

32. Brannigan, J. A. & Wilkinson, A. J. Protein engineering 20 years on. *Natural reviews. Molecular cell biology* **3**, 964-70 (2002).

33. Breiter, D. R., Resnik, E. & Banaszak, L. J. Engineering the quaternary structure of an enzyme: construction and analysis of a monomeric form of malate dehydrogenase from *Escherichia coli*. *Protein science* **3**, 2023-32 (1994).

34. Sieber, C., Ploger, F., Schwappacher, R., Bechtold, R., Hanke, M., Kawai, S., Muraki, Y., Katsuura, M., Kimura, M., Rechtman, M. M., Henis, Y. I., Pohl, J. & Knaus, P. Monomeric and dimeric GDF-5 show equal type I receptor binding and oligomerization capability and have the same biological activity. *Biological chemistry* **387**, 451-60 (2006).

35. Koshkin, A., Knudsen, G. M. & Ortiz De Montellano, P. R. Intermolecular interactions in the AhpC/AhpD antioxidant defense system of *Mycobacterium tuberculosis*. *Archives of biochemistry and biophysics* **427**, 41-7 (2004).

36. Chen, Y., Birck, C., Samama, J. P. & Hulett, F. M. Residue R113 is essential for PhoP dimerization and function: a residue buried in the asymmetric PhoP dimer interface determined in the PhoPN three-dimensional crystal structure. *Journal of bacteriology* **185**, 262-73 (2003).

37. Porvari, K. S., Herrala, A. M., Kurkela, R. M., Taavitsainen, P. A., Lindqvist, Y., Schneider, G. & Vihko, P. T. Site-directed mutagenesis of prostatic acid phosphatase. Catalytically important aspartic acid 258, substrate specificity, and oligomerization. *The Journal of biological chemistry* **269**, 22642-6 (1994).

38. Ottosen, M. B., Bjornberg, O., Norager, S., Larsen, S., Palfey, B. A. & Jensen, K. F. The dimeric dihydroorotate dehydrogenase A from *Lactococcus lactis* dissociates reversibly into inactive monomers. *Protein science* **11**, 2575-83 (2002).

39. Franke, I., Meiss, G., Blecher, D., Gimadutdinow, O., Urbanke, C. & Pingoud, A. Genetic engineering, production and characterisation of monomeric variants of the dimeric *Serratia marcescens* endonuclease. *FEBS Letters* **425**, 517-22 (1998).

40. Dai, H., Mao, Q., Yang, H., Huang, S. & Chang, Z. Probing the roles of the only universally conserved leucine residue (Leu122) in the oligomerization and chaperone-like activity of

Mycobacterium tuberculosis small heat shock protein Hsp16.3. *Journal of protein chemistry* **19**, 319-26 (2000).

41. Jackson, R. M., Gelpi, J. L., Cortes, A., Emery, D. C., Wilks, H. M., Moreton, K. M., Halsall, D. J., Sleigh, R. N., Behan-Martin, M., Jones, G. R., Clarke, A. R. & Holbrook, J. J. Construction of a stable dimer of *Bacillus stearothermophilus* lactate dehydrogenase. *Biochemistry* **31**, 8307-14 (1992).

42. Bisgrove, B. W., Andrews, M. E. & Raff, R. A. Fibropellins, products of an EGF repeat-containing gene, form a unique extracellular matrix structure that surrounds the sea urchin embryo. *Developmental biology* **146**, 89-99 (1991).

43. Hunt, L. T. & Barker, W. C. Avidin-like domain in an epidermal growth factor homolog from a sea urchin. *The FASEB journal* **3**, 1760-4 (1989).

44. Laitinen, O. H., Nordlund, H. R., Hytonen, V. P., Uotila, S. T., Marttila, A. T., Savolainen, J., Airenne, K. J., Livnah, O., Bayer, E. A., Wilchek, M. & Kulomaa, M. S. Rational design of an active avidin monomer. *The Journal of biological chemistry* **278**, 4010-4 (2003).

45. Laitinen, O. H., Airenne, K. J., Marttila, A. T., Kulik, T., Porkka, E., Bayer, E. A., Wilchek, M. & Kulomaa, M. S. Mutation of a critical tryptophan to lysine in avidin or streptavidin may explain why sea urchin fibropellin adopts an avidin-like domain. *FEBS Letters* **461**, 52-8 (1999).

46. Boersma, M. D., Sadowsky, J. D., Tomita, Y. A. & Gellman, S. H. Hydrophile scanning as a complement to alanine scanning for exploring and manipulating protein-protein recognition: application to the Bim BH3 domain. *Protein science* **17**, 1232-40 (2008).

47. Cunningham, B. C. & Wells, J. A. High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. *Science* **244**, 1081-5 (1989).

48. Clackson, T. & Wells, J. A. A hot spot of binding energy in a hormone-receptor interface. *Science* **267**, 383-6 (1995).

49. Zhu, H., Smith, P., Wang, L. K. & Shuman, S. Structure-function analysis of the 3' phosphatase component of T4 polynucleotide kinase/phosphatase. *Virology* **366**, 126-36 (2007).

50. Montioli, R., Cellini, B., Bertoldi, M., Paiardini, A. & Voltattorni, C. B. An engineered folded PLP-bound monomer of *Treponema denticola* cystalysin reveals the effect of the dimeric structure on the catalytic properties of the enzyme. *Proteins* **74**, 304-17 (2009).

51. Pradhan, A., Mukherjee, P., Tripathi, A. K., Avery, M. A., Walker, L. A. & Tekwani, B. L. Analysis of quaternary structure of a [LDH-like] malate dehydrogenase of *Plasmodium falciparum* with oligomeric mutants. *Molecular and cellular biochemistry* **325**, 141-8 (2009).

52. Salminen, A., Efimova, I. S., Parfenyev, A. N., Magretova, N. N., Mikalahti, K., Goldman, A., Baykov, A. A. & Lahti, R. Reciprocal effects of substitutions at the subunit interfaces in hexameric pyrophosphatase of *Escherichia coli*. Dimeric and monomeric forms of the enzyme. *The Journal of biological chemistry* **274**, 33898-904 (1999).

53. DeLano, W. L. The PyMOL molecular graphics system. (DeLano Scientific, San Carlos, CA, USA, 2006).

54. Neshich, G., Rocchia, W., Mancini, A. L., Yamagishi, M. E., Kuser, P. R., Fileto, R., Baudet, C., Pinto, I. P., Montagner, A. J., Palandrani, J. F., Krauchenco, J. N., Torres, R. C., Souza, S.,

Togawa, R. C. & Higa, R. H. JavaProtein Dossier: a novel web-based data visualization tool for comprehensive analysis of protein structure. *Nucleic acids research* **32**, 595-601 (2004).

55. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *Journal of molecular biology* **372**, 774-97 (2007).

56. Tomovic, A. & Oakeley, E. J. Computational structural analysis: multiple proteins bound to DNA. *PLoS one* **3**, e3243 (2008).

57. Dobson, R. C., Griffin, M. D., Jameson, G. B. & Gerrard, J. A. The crystal structures of native and (*S*)-lysine-bound dihydrodipicolinate synthase from *Escherichia coli* with improved resolution show new features of biological significance. *Acta crystallographica. Section D, Biological crystallography* **61**, 1116-24 (2005).

58. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research* **22**, 4673-80 (1994).

59. Stratagen. Quikchange site-directed mutagenesis kit: instruction manual. (2004).

60. Sanger, F., Nicklen, S. & Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **74**, 5463-7 (1977).

61. Fenton, W. A. & Horwich, A. L. GroEL-mediated protein folding. *Protein science: a publication of the protein society* **6**, 743-60 (1997).

62. Goloubinoff, P., Gatenby, A. A. & Lorimer, G. H. GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose bisphosphate carboxylase oligomers in *Escherichia coli*. *Nature* **337**, 44-7 (1989).

63. Schagger, H. Blue-native gels to isolate protein complexes from mitochondria. *Methods in cell biology* **65**, 231-44 (2001).

64. Bulheller, B. M., Rodger, A. & Hirst, J. D. Circular and linear dichroism of proteins. *Physical chemistry chemical physics* **9**, 2020-35 (2007).

65. Sreerama, N., Venyaminov, S. Y. & Woody, R. W. Analysis of protein circular dichroism spectra based on the tertiary structure classification. *Analytical biochemistry* **299**, 271-4 (2001).

66. Pelton, J. T. & McLean, L. R. Spectroscopic methods for analysis of protein secondary structure. *Analytical biochemistry* **277**, 167-76 (2000).

67. Kelly, S. M. & Price, N. C. The use of circular dichroism in the investigation of protein structure and function. *Current protein & peptide science* **1**, 349-84 (2000).

68. Provencher, S. W. & Glockner, J. Estimation of globular protein secondary structure from circular dichroism. *Biochemistry* **20**, 33-7 (1981).

69. van Stokkum, I. H., Spoelder, H. J., Bloemendal, M., van Grondelle, R. & Groen, F. C. Estimation of protein secondary structure and error analysis from circular dichroism spectra. *Analytical biochemistry* **191**, 110-8 (1990).

70. Sreerama, N. & Woody, R. W. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Analytical biochemistry* **287**, 252-60 (2000).

71. Pearce, F. G., Dobson, R. C., Weber, A., Lane, L. A., McCammon, M. G., Squire, M. A., Perugini, M. A., Jameson, G. B., Robinson, C. V. & Gerrard, J. A. Mutating the tight-dimer interface of dihydrodipicolinate synthase disrupts the enzyme quaternary structure: toward a monomeric enzyme. *Biochemistry* **47**, 12108-17 (2008).

72. Dobson, R. C., Devenish, S. R., Turner, L. A., Clifford, V. R., Pearce, F. G., Jameson, G. B. & Gerrard, J. A. Role of arginine 138 in the catalysis and regulation of *Escherichia coli* dihydrodipicolinate synthase. *Biochemistry* **44**, 13007-13 (2005).

Chapter Four

Purification & characterization of interface mutant DHDPS-A204R from *M. tuberculosis*

4.1 Introduction

A key goal in this work was to disrupt the weak interface of the tetrameric protein, *M. tuberculosis* DHDPS, and create and characterize discrete, stable, folded dimeric units. As presented in the previous chapter, BN-PAGE showed the tetrameric structure to be disrupted in two interface mutants A204D and A204R *M. tuberculosis* DHDPS, in which alanine 204 was replaced with aspartate and arginine, respectively. The most promising interface mutant according to BN-PAGE was DHDPS-A204R, as it seemed to exist predominantly as a dimer. DHDPS-A204R, was therefore chosen for detailed characterization and comparison to wild-type *M. tuberculosis* DHDPS. Of particular interest were changes in the catalytic ability in comparison to the wild-type enzyme. Excitingly, the solid-state structure was determined by X-ray crystallography as part of a collaborative project with Dr Manfred Weiss and Linda Schuldt, from the EMBL Hamburg Outstation. This allowed differences between mutant A204R and wild-type *M. tuberculosis* DHDPS to be understood in the context of structural data, thus clarifying whether these differences were due to changes at the active site propagated by the mutation of alanine to arginine or to the change in oligomeric state.

4.2 Enzyme over-expression & purification

The rigorous characterization of A204R *M. tuberculosis* DHDPS required production and purification of milligram quantities of protein. As discussed in chapter two, as a recombinant protein in *E. coli*, the majority of *M. tuberculosis* DHDPS is expressed in insoluble form, resulting in a final yield of ~0.7 mg per litre of LB medium using an optimized procedure. Similarly, the majority of *M. tuberculosis* DHDPS-A204R was expressed as an insoluble protein in *E. coli* **(Figure 4.1, lane 4)**; however, an even lower yield of 0.4 mg per litre was observed in small scale preparations of DHDPS-A204R (in chapter three, section 3.6.2). This is not surprising given that disruption of the weak interface results in exposure of previously buried hydrophobic surfaces and therefore is likely to cause decreased enzyme solubility and stability.

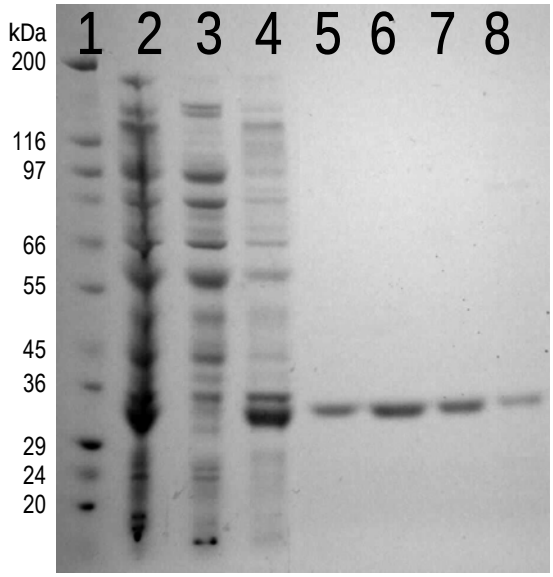


Figure 4.1: Small scale purification of *M. tuberculosis* DHDPS-A204R, over-expressed in *E. coli* BL21 (DE3) pGroESL pMTBA204R, examined using SDS-PAGE. The MW marker, in lane 1, enabled identification of His_6-tagged DHDPS-A204R as the 34 kDa band. Protein suspension from various stages of purification was loaded into other lanes as follows: 2 - crude extract, 3 - supernatant after centrifugation, 4 - re-suspended cell pellet from centrifugation, 5 & 6 - eluted from His_6-tag affinity column, 7 - dialyzed into pyruvate containing buffer, 8 - dialyzed out of pyruvate. Lanes 3 and 4 contain the soluble and insoluble protein fractions, respectively. The decreased band intensity in lane 8, suggests loss of protein in buffer lacking pyruvate, possibly by the formation of aggregates.

4.2.1 The inclusion of pyruvate in storage buffers

The stabilizing effect of pyruvate on wild-type *M. tuberculosis* DHDPS had previously been noted (in chapter two, section 2.3.2), and consequently had been included in all purification buffers except the final storage buffer of 20 mM Tris.HCl, pH 8.0, 250 mM NaCl, 5 % glycerol, and 2 mM β-mercaptoethanol (β-ME). During the purification of A204R *M. tuberculosis* DHDPS, dialyzing the enzyme into a buffer without pyruvate seemed to result in a loss of protein as observed by SDS-PAGE **(Figure 4.1, lane 8)**, and of activity (Table 3.3, chapter three, section 3.6.2).

The thermal stability of DHDPS-A204R was initially investigated using differential scanning fluorescence (DSF), an assay where a sharp increase in fluorescence **(**shown in **Figure 4.2A)**

corresponds to the unfolding temperature, or apparent melting temperature, T_m^{app}.[1] In storage buffer containing pyruvate, interface mutant DHDPS-A204R had a T_m^{app} of 51.6 (±0.1) ºC, compared with 85.3 (±0.1) ºC determined for wild-type *M. tuberculosis* DHDPS. The decreased thermal stability of DHDPS-A204R suggested that the interface mutation had successfully disrupted quaternary structure, as a decrease in unfolding temperature typically reflects either decreased structural order, increased flexibility or protein misfolding,[1] and the CD spectra showed DHDPS-A204R to be folded (Figure 3.9, chapter three, section 3.7.2).

DSF was also used to quantify the stabilizing effect of pyruvate at two enzyme concentrations in a pH 8.0 buffer containing Tris and NaCl at the same concentration as the buffer commonly used in AUC analysis **(Table 4.1)**. Increasing concentrations of pyruvate were observed to have a stabilising affect on T_m^{app} **(Figure 4.2B)**. A similar pyruvate effect has been observed in other DHDPS enzymes (Dr Grant Pearce, *pers. comm.*), and is assumed to be due to pyruvate introducing rigidity to the active site.

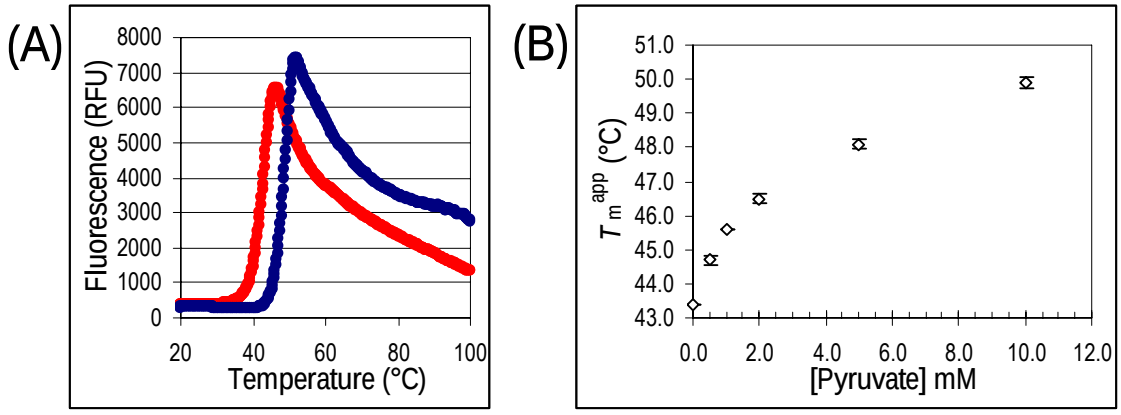


Figure 4.2: Determination of T_m^{app} using differential scanning fluorescence. (A) The melting curve for 0.3 mg.mL^{-1} of DHDPS-A204R in buffer of 20 mM Tris.HCl, pH 8.0, 150 mM NaCl, 10 mM pyruvate (blue), and in the absence of pyruvate (red). (B) An increase in T_m^{app} with increasing pyruvate concentrations (in 20 mM Tris.HCl, 150 mM NaCl, pH 8.0) can be observed.

Table 4.1: The T_m^{app} of DHDPS-A204R in pH 8.0 buffer composed of 20 mM Tris.HCl, 150 mM NaCl and varying pyruvate concentrations, as determined by DSF.

[Pyruvate] mM	Apparent melting temperature (ºC)	
	for 0.3 mg.mL^{-1} DHDPS-A204R	for 0.7 mg.mL^{-1} DHDPS-A204R
0.0	43.4 (±0.0)	43.4 (±0.0)
0.5	44.7 (±0.1)	44.6 (±0.3)
1.0	45.6 (±0.0)	47.0 (±0.0)
2.0	46.5 (±0.1)	47.6 (±0.0)
5.0	48.1 (±0.1)	49.4 (±0.0)
10.0	49.9 (±0.1)	50.6 (±0.0)

A shift of ~7 ºC in T_m^{app} as a result of the inclusion of 10 mM pyruvate can be observed at both concentrations of A204R *M. tuberculosis* DHDPS **(Table 4.1)** suggesting that the effect of pyruvate is not strongly dependent on protein concentration. A comparable shift of 7.7 (±0.3) ºC in T_m^{app} was observed for wild-type *M. tuberculosis* DHDPS (chapter two, section 2.3.2), which indicates that the pyruvate stabilizing effect is not specific for DHDPS-A204R. The instability of the mutant compared to wild-type resulted in the inclusion of pyruvate in all buffers including the final storage buffer for DHDPS-A204R.

4.2.2 Large scale purification of A204R *M. tuberculosis* DHDPS

The optimized protocol, as described in chapter two, section 2.3.3, was further improved by the use of auto-inducing media, which resulted in high density cultures.[2] Auto-inducing media uses glucose to suppress protein expression from the pET protein expression system **(**used in the plasmid construct pMTB02 and pMTBA204R, details in **Appendix A)** until after the culture has grown to saturation.[2] Several batches of auto-induced culture were harvested and cell pellets frozen, allowing for a large scale preparation of *M. tuberculosis* DHDPS-A204R from 67.4 g (wet weight) of *E. coli* BL21 (DE3) pGroESL pMTBA204R cells **(Table 4.2)**. As with previous purifications, crude cell-free lysate was produced by ultrasonication and the insoluble proteins removed by centrifugation **(Figure 4.3, lane 2 and 3**, respectively**)**. Elution from the His_6-tag affinity column resulted in a 57-fold increase in protein purity **(Table 4.2).** Subsequent to incubation with TEV protease, the cleaved DHDPS-A204R was passed through a His_6-tag affinity column to remove the His_6-tag peptide fragment **(Figure 4.3, lane 5)** and similar to wild-type, an increase in total activity was seen after His_6-tag cleavage **(Table 4.2)**. Minor impurities were apparent upon concentration (with an ultrafiltration spin column), and

elution from a gel-filtration column yielded 9 mg of protein that was homogeneous as judged by SDS-PAGE **(Figure 4.3, lane 6)**.

Table 4.2: Large scale purification of *M. tuberculosis* DHDPS-A204R from *E. coli* BL21 (DE3) pGroESL pMTBA204R (67.4 g of cells).

Purification step	Volume (mL)	Protein[a] (mg)	Total activity[b] (units[b])	Specific activity (units[b]/mg)	Relative total activity	Purification
Centrifuged	150	3281	22.4	0.0068	-	1-fold
Affinity (His_6-tagged)	28.0	23	8.8	0.39	40 %	57-fold
Affinity (Non-tagged)	45.0	13	13.8	1.05	62 %	154-fold
Gel filtration	17.5	9	9.7	1.12	43 %	164-fold

[a] Protein concentration was determined using the Bradford assay, detailed in chapter 6, section 6.3.1.

[b] Enzyme activity was determined using the coupled assay, described in chapter 2, section 2.2.3 and 1 unit was defined as the consumption of 1 µmol of NADPH per second.

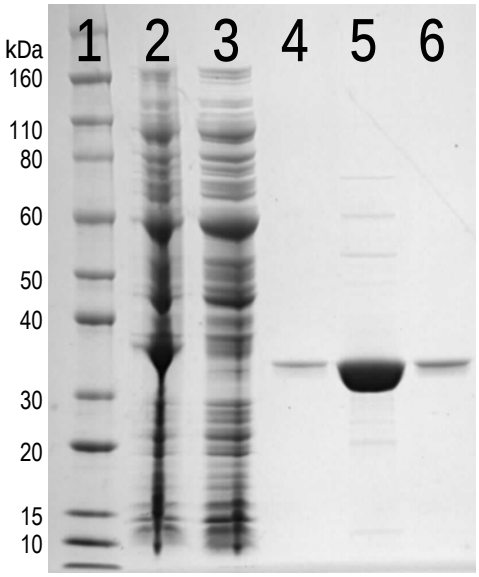


Figure 4.3: Typical large scale purification for *M. tuberculosis* DHDPS-A204R, over-expressed in *E. coli* BL21 (DE3) pGroESL pMTBA204R, examined using denaturing SDS-PAGE. The MW marker, in lane 1, enabled identification of non-His_6-tagged *M. tuberculosis* DHDPS-A204R as the 31 kDa band. Preparations from purification steps were loaded into other lanes as follows: 2 - crude extract, 3 - supernatant after centrifugation, 4 - after His_6-tag cleavage by TEV protease and affinity chromatography, 5 - concentrated *via* an ultrafiltration spin column, 6 - eluted from gel-filtration column.

4.3 Biochemical characterization of DHDPS-A204R

Biochemical characterization of A204R *M. tuberculosis* DHDPS was carried out to determine the effect of the mutation and associated disrupted quaternary structure on enzyme function. The coupled assay was used to measure initial-rates by coupling DHDPS activity to the activity of DHDPR (chapter two, section 2.2.3). Determination of kinetic parameters in combination with the investigation of factors that affect catalytic activity, such as pH, ionic strength, and temperature, allowed a comprehensive comparison between wild-type and A204R *M. tuberculosis* DHDPS. Unless otherwise stated, the enzyme used was in storage buffer, which includes 10 mM pyruvate. For the kinetic analysis the extra pyruvate added with the enzyme into the assay mixture was taken into account (chapter six, section 6.9.4).

4.3.1 The effect of pH, buffer and ionic strength on enzyme activity

The pH optimum was determined in the manner outlined in chapter two, section 2.4.1, with the coupled assay system and using a series of buffers covering a pH range of 6 to 9 adjusted to the same ionic strength (IS) by addition of NaCl. The highest activity measurements for DHDPS-A204R of 1.02 (±0.10) and 0.97 (±0.11) $\mu mol.s^{-1}.mg^{-1}$ were recorded at pH 7.75 and 8.25 **(Figure 4.4A)**, showing a similar trend to that observed in wild-type *M. tuberculosis* DHDPS (chapter two, section 2.4.1). These values were close to those measured for wild-type enzyme at pH 7.75 and 8.25 of 1.34 (±0.05) and 1.39 (±0.04) $\mu mol.s^{-1}.mg^{-1}$. Therefore, it seemed that the alanine to arginine mutation has not changed the optimal pH range of 7.75 to 8.25, nor had the overall activity been grossly affected; however, steady-state kinetics were needed for detailed comparison.

Unlike the wild-type enzyme, the type of buffer used seemed to have a dramatic effect on the catalytic activity of DHDPS-A204R, with an almost two-fold decrease observed at pH 8.0 between enzyme buffered with bicine compared to HEPES. All of the buffers used were zwitterionic buffers, as recommended by Good.[3] Thus, it would seem that the activity of the interface mutant is more sensitive to differences in buffer components than the wild-type *M. tuberculosis* DHDPS, perhaps reflecting solubility issues.

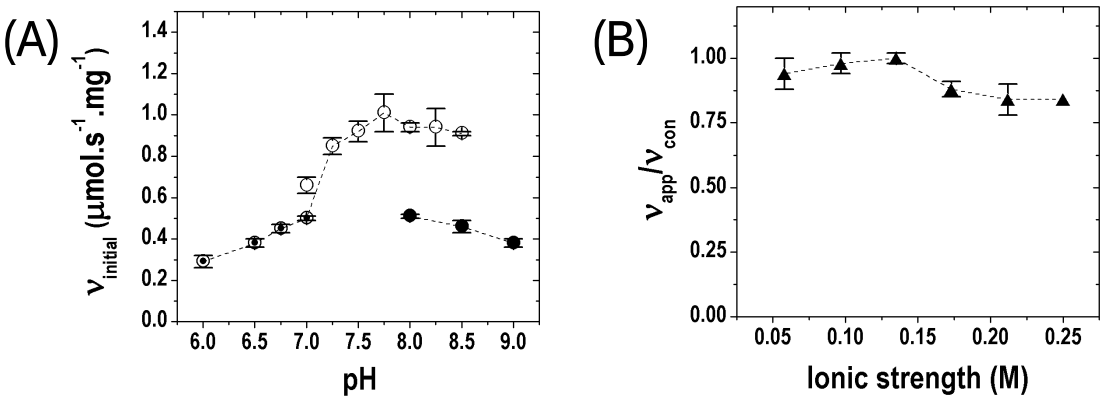


Figure 4.4: The optimum conditions for DHDPS-A204R activity determined for varying (A) pH with different buffers (◉ MES, ○ HEPES, ● Bicine), and (B) ionic strength with salt concentration. The data in the second plot are normalized, the apparent rate (v_{app}), with units AU/s, was divided by highest measured rate (v_{con}) of enzyme activity. Each data point was measured in at least triplicate, the error bars show the standard deviation, and the dashed line show the general trend.

The effect of ionic strength on the enzymatic reaction was examined using a method analogous to that described in chapter two, section 2.4.2, with a series of 20 mM HEPES pH 8.25 buffers adjusted with NaCl to cover a range of ionic strengths. There was no significant difference between the measurements below 0.15 M or above 0.2 M, but a small decrease in activity was observed in the physiological range of 0.15-0.2 M **(Figure 4.4B)**. Ionic strengths above 0.2 M showed a clear decrease in activity for wild-type *M. tuberculosis* DHDPS, as is typical for most enzymes,[4] and 0.15-0.2 M was determined as optimal for its activity. For DHDPS-A204R, low ionic strength (0.05-0.15 M) seems to be optimal for activity; however, the effect of ionic strength is small. Given these observations with ionic strength and the optimal pH range of 7.75 to 8.25, the same buffer stock used for wild-type, 200 mM HEPES, 50 mM NaCl, pH 8.0, was prepared for kinetic analysis of the interface mutant, which gives the assay mixture an overall ionic strength of ~0.18 M **(Appendix D)**.

4.3.2 Thermal stability

To investigate the thermal stability of enzyme activity, DHDPS-A204R was incubated for five minutes at various temperatures before measuring initial-rates with the coupled assay at 30 ºC. Wild-type and A204R *M. tuberculosis* DHDPS were incubated in either storage buffer containing 10 mM pyruvate or buffer without pyruvate. For DHDPS-A204R, apparent melting temperatures (T_m^{app}) of 47 (±1) ºC and 53 (±1) ºC were determined in the absence and

presence of pyruvate, respectively **(Figure 4.5)**. The T_m^{app} in storage buffer containing 10 mM pyruvate is within error of the 51.6 (±0.1) ºC determined by DSF (section 4.2.1). The 6 ºC difference in T_m^{app} caused by 10 mM pyruvate was similar to the ~7 ºC change observed using DSF (Table 4.1, section 4.2.1). Parallel experiments with wild-type *M. tuberculosis* DHDPS also showed an increase in T_m^{app} in the presence of pyruvate, to 86 (±1) ºC from 77 (±1) ºC **(Figure 4.5)**, confirming that the pyruvate stabilizing effect is not specific to the interface mutant.

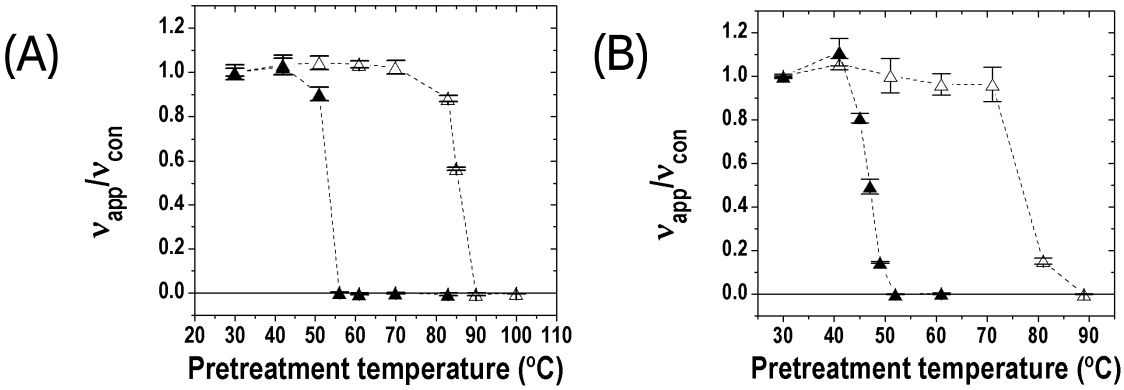


Figure 4.5: **The thermal stability of interface mutant A204R (▲) was compared to wild-type *M. tuberculosis* DHDPS (△) in (A) the storage buffer containing pyruvate and (B) buffer without pyruvate [20 mM Tris.HCl, pH 8.0, 250 mM NaCl, 5 % glycerol, 2 mM β-ME]. The enzyme activity was determined using the coupled assay after pre-treatment by a 5 minute incubation at the indicated temperature. The data are normalized by dividing the apparent rate (ν_{app}) by the rate after incubation at the lowest temperature measured (ν_{con}). Each data point was measured in triplicate, the error bars show the standard deviation, and the dashed line shows the general trend.**

While the effect of pyruvate on T_m^{app} was similar between the wild-type and mutant enzyme, significant differences were observed during thermal melts. While, the activity of DHDPS-A204R was lower than that of wild-type DHDPS at all incubation temperatures, the removal of pyruvate via buffer exchange substantially lowered the activity of DHDPS-A204R, decreasing it from 0.89 (±0.04) to 0.156 (±0.005) $\mu mol.s^{-1}.mg^{-1}$. Thus, in the absence of pyruvate, DHDPS-A204R had approximately 18 % of its original activity and 8 % that of the wild-type enzyme. This is consistent with earlier observations during purification (section 4.2) that suggest DHDPS-A204R becomes unstable in the absence of pyruvate.

In these thermal stability measurements, data were not normalized to the rate after incubation for 5 minutes at 30 ºC. This is in contrast to previous melt experiments (chapter two, section 2.4.3) which were normalized relative to the rate of enzyme without heat treatment. Incubation for 5 minutes at 30 ºC had no effect on wild-type DHDPS whereas the catalytic activity of DHDPS-A204R was significantly increased (compared with zero incubation time), both in the presence and absence of pyruvate, from 0.89 (±0.04) to 1.39 (±0.03) $\mu mol.s^{-1}.mg^{-1}$ and from 0.156 (±0.005) to 0.528 (±0.002) $\mu mol.s^{-1}.mg^{-1}$, respectively. Therefore, the data presented in **Figure 4.5** were normalized using the rate measured after incubation at the lowest temperature. This apparent heat activation of DHDPS-A204R is hard to rationalize. Heat activation was also observed for the interface mutants of *E. coli* DHDPS; however, this was caused by the presence of an pyruvate analogue, α-ketoglutarate, and its subsequent removal by incubation with pyruvate.[5] For DHDPS-A204R, the presence of pyruvate in all purification buffers would have prevented binding of pyruvate analogues, such as α-ketoglutarate. It seems more likely that the heat activation observed for DHDPS-A204R is due to dissociation of protein aggregates in the incubation buffer. However, further experiments would be useful to confirm this.

4.3.3 Steady-state kinetics of A204R *M. tuberculosis* DHDPS

Steady-state kinetics describes the binding of substrates and release of products in terms of mathematical models, such as the Michaelis-Menten outlined in chapter two, section 2.4.4. The calculations are simplified by assuming a "steady-state", where the reaction's enzyme-bound intermediates are being consumed as quickly as they are produced; that is the initial-rate before it is affected by product accumulation or substrate consumption.

Initial-rate data for kinetic characterization of A204R *M. tuberculosis* DHDPS were determined similarly to that outlined for wild-type enzyme, using a five-by-five matrix, varying pyruvate and (*S*)-ASA concentrations from 0.2 to 5 times their apparent K_M values (which were determined to be 0.20 and 1.1 mM for pyruvate and (*S*)-ASA, respectively). The initial-rate of product formation was measured in at least triplicate, and was reproducible (±10 %). Data were fitted with the same models used for wild-type enzyme analysis, described in detail in chapter two, section 2.4.4, using the program ENZFITTER.

The ping-pong (*S*)-ASA substrate inhibition model provided the best fit, as shown by its R^2 value of 0.99, and yielded K_M constants of 0.29 (±0.03) mM for pyruvate and 1.4 (±0.2) mM for (*S*)-ASA. Whether substrate inhibition with (*S*)-ASA is observed in DHDPS is debated in the literature **(**details in **Appendix G)**.[6-8] However, the determined $K_i^{(S)\text{-ASA}}$ value of 10 (±4) mM shows low precision due to the value lying outside the range of (*S*)-ASA concentrations (5.5 to 0.22 mM) used in this study; consequently, the ping-pong model without substrate inhibition was used to fit these data giving similar K_M constants of 0.33 (±0.03) mM for pyruvate and 1.1 (±0.1) mM for (*S*)-ASA. The phenomenon of substrate inhibition with DHDPS from *M. tuberculosis* is discussed in more detail in **Appendix G**.

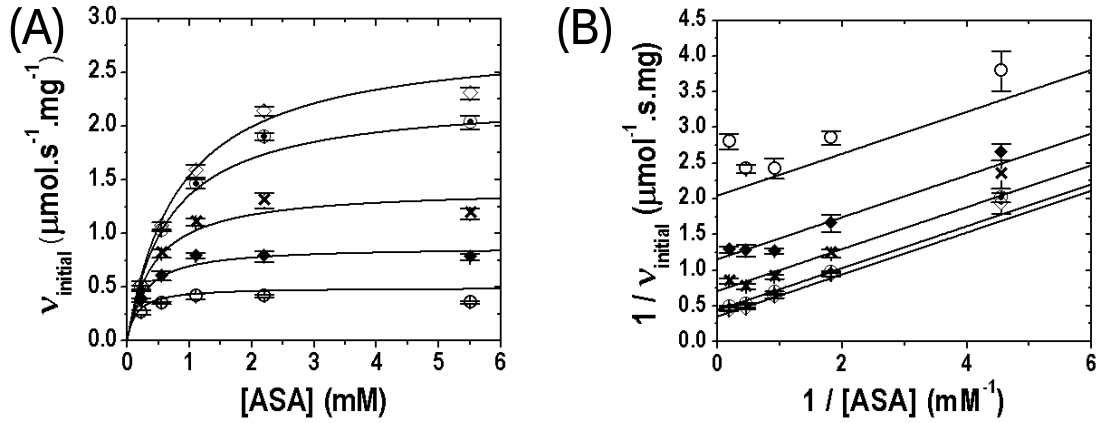


Figure 4.6: **(A) The initial velocity of A204R *M. tuberculosis* DHDPS was measured with the coupled assay (buffered at pH 8.0 with 20 mM HEPES, 50 mM NaCl) over varying concentrations of both substrates, (*S*)-ASA and pyruvate, shown on the x-axis and with symbols (◇ 1.0 mM, ◉ 0.50 mM, ✕ 0.20 mM, ◆ 0.10 mM, ○ 0.050 mM pyruvate), respectively. Each point was measured in triplicate and the data were fitted with the ping-pong model, giving an R^2 of 0.98, using the program ENZFITTER. (B) Reciprocal values were plotted to give a Lineweaver-Burk transformation, and compared with the parallel lines predicted by the fitted model.**

The closeness of data points to the parallel lines in the Lineweaver-Burk plot **(Figure 4.6B, 4.7A)** indicates that substrates bind to A204R *M. tuberculosis* DHDPS in a ping-pong fashion. The upward deviation of points near the y-axis **(Figure 4.6B)** is diagnostic of inhibition at high (*S*)-ASA concentrations.[9,10] Kinetic analysis of wild-type *M. tuberculosis* DHDPS showed no substrate inhibition; although, this is likely due to the low (*S*)-ASA concentrations

being investigated **(**further comments in **Appendix G)**. The compulsory ordered nature of the ping-pong mechanism suggests that (*S*)-ASA could inhibit activity at high concentrations by binding prior to pyruvate and forming a dead-end complex.[11]

A value for *V* of 3.8 (±0.2) µmol.s^{-1}.mg^{-1} was determined for DHDPS-A204R, which corresponds to a catalytic turnover number, k_{cat}, of 119 s^{-1} per active site. The kinetic parameters are summarized and compared with wild-type *M. tuberculosis* DHDPS in **Table 4.3**, including k_{cat}/K_M, which describes catalytic efficiency.

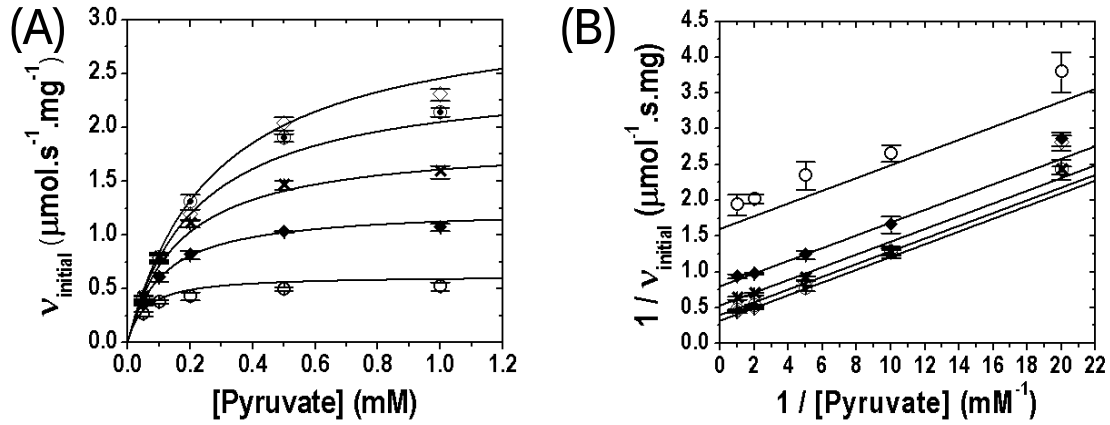


Figure 4.7: The initial velocity of A204R *M. tuberculosis* DHDPS was measured with varying concentrations of both substrates, pyruvate and (*S*)-ASA, shown on the x-axis and with symbols (◇ 5.5 mM, ◉ 2.2 mM, ×1.1 mM, ◆ 0.55 mM, ○ 0.22 mM, (*S*)-ASA), respectively. Each point was measured at least in triplicate and the data were fitted with the ping-pong model, giving an R^2 of 0.98, using the program ENZFITTER. (B) Reciprocal values of the raw data were plotted to give a Lineweaver-Burk transformation.

Investigations with dimeric mutants of *E. coli* DHDPS saw a decrease in k_{cat} to 1.4 % and 2.5 % that of the wild-type tetrameric enzyme, for DHDPS-L197Y and DHDPS-L197D, respectively.[5] In contrast, the k_{cat} and *V* values determined for the interface mutant DHDPS-A204R are very similar to the values determined for wild-type *M. tuberculosis* DHDPS **(Table 4.3)**. However, K_M values for both substrates are considerably larger than those determined for the wild-type enzyme. This increase in K_M of 2-fold for pyruvate and 3-fold for (*S*)-ASA **(Table 4.3)** indicates that the binding of the substrates is not optimal in DHDPS-A204R. Consequently it seems that the mutation of alanine to arginine has decreased the catalytic efficiency (k_{cat}/K_M) of *M. tuberculosis* DHDPS, although the maximum rate of catalysis has not been affected, unlike previous experiments with *E. coli* DHDPS mutants.

Table 4.3: Kinetic parameters determined by fitting rate measurements.

	Wild-type DHDPS	DHDPS-A204R	Comparison
Kinetic model	Ping-pong	Ping-pong	
K_M for (*S*)-ASA (mM)	0.43 (±0.02)	1.1 (±0.1)	3-fold increase
K_M for pyruvate (mM)	0.17 (±0.01)	0.33 (±0.03)	2-fold increase
V (μmol.s^{-1}.mg^{-1})	4.42 (±0.08)	3.8 (±0.2)	Slight decrease
Molecular Weight (Da)	31156.5	31241.6	
e_0 (mg of enzyme per μmol)	31.16	31.24	
k_{cat} (s^{-1})	132 (±3)	119 (±6)	Slight decrease
k_{cat}/K_M for (*S*)-ASA (s^{-1}.mM^{-1})	300 (±20)	110 (±20)	3-fold decrease
k_{cat}/K_M for pyruvate (s^{-1}.mM^{-1})	780 (±60)	360 (±50)	2-fold decrease

4.4 Determining the quaternary structure of DHDPS-A204R

The unexpected similarity between the k_{cat} of DHDPS-A204R and the wild-type enzyme could reflect that the tetrameric quaternary structure has not been disrupted in the mutant, as was indicated with BN-PAGE (chapter three, section 3.6.3). To confirm the quaternary structure(s) adopted by DHDPS-A204R in solution, two biophysical techniques, gel-filtration chromatography and analytical ultracentrifugation (AUC), were used.

4.4.1 Determining quaternary structure in solution

Initially, the oligomeric state of DHDPS-A204R was analyzed using sedimentation velocity (SV) experiments performed at 40000 rpm, applying a large centrifugal force. The resulting concentration boundary and rate of sedimentation were determined by radial scans at 230 or 280 nm every 8 minutes without averaging. Radial scans were analyzed using the program SEDFIT, by fitting the continuous sedimentation coefficient [*c(s)*] distribution model.[12] The partial specific volume ($\bar{v}$) of DHDPS-A204R was calculated to be 0.7400 mL.g^{-1} using the program SEDNTERP.[13] As with the wild-type enzyme, AUC experiments were performed at 20 ºC in 20 mM Tris.HCl, 150 mM NaCl, pH 8.0, which is a commonly used buffer, thus its viscosity and density are well known, and easily calculated with SEDNTERP.

Data collected at a low concentration of 0.06 mg.mL^{-1} DHDPS-A204R and fitted with a [*c(s)*] distribution, showed one main peak of 2.7 S (96 % of signal), with a much smaller secondary peak of 5.7 S **(Figure 4.8B)**. Plots of raw data overlaid with the calculated fits were produced for visual inspection of the analysis **(Figure 4.8A)**. The random distribution of the residuals and low r.m.s.d. of 0.0037 indicate the quality of the fit. The sedimentation coefficients

observed (s^*) were converted into standardized sedimentation coefficients ($s°_{20,w}$) of 2.8 and 5.9 S, using the program SEDNTERP.[13]

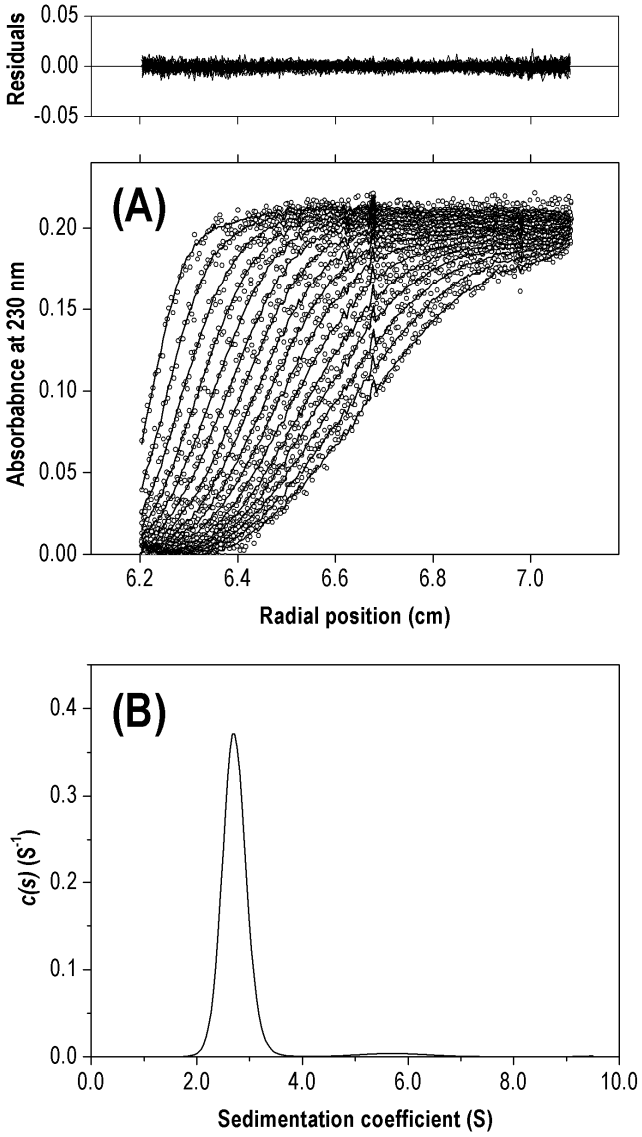


Figure 4.8: **SV analysis of DHDPS-A204R (0.06 mg.mL^{-1}) in 20 mM Tris.HCl, 150 mM NaCl, pH 8.0 at 20 ºC. (A) Absorbance at 230 nm (○) is plotted at time intervals of 16 minutes and overlaid with the nonlinear least-squares fit (solid line) to a continuous sedimentation coefficient [*c(s)*] distribution model.[12] (B) The [*c(s)*] distribution is plotted as a function of s^* (in units of Svedberg, S). The fit was obtained using a resolution of 200 species with $\bar{\nu}$ = 0.7400 mL.g^{-1}, ρ = 1.005 g.mL^{-1}, η = 1.021 cP and f/f_0 = 1.22904. The r.m.s.d. and Z-test for the fit were 0.0037 and 0.69, respectively. A radial-dependent, time-invariant (TI) baseline was subtracted from data to account for optical imperfections.**

The sedimentation coefficient values for different quaternary structures were calculated by hydrodynamic modelling, using the program HYDROPRO.[14] The hydrodynamic models for different oligomers were based on the X-ray crystal structure for wild-type *M. tuberculosis* DHDPS (PDB entry 1XXX). HYDROPRO predicted the sedimentation coefficients (*s*) for the monomer, dimer and tetramer in the experimental buffer to be 2.8 S, 4.4 S and 6.7 S, respectively. This suggests that the primary oligomeric species (s^* = 2.7 S) observed for DHDPS-A204R is a monomer under the conditions of this experiment, rather than the expected dimer. The secondary peak falls between the predicted values for dimer and tetramer, possibly suggesting the presence of higher order oligomers. However, there is only a trace amount of this oligomeric species, and it may simply be an artefact of analysis.

Using SEDFIT, the [*c(s)*] distribution was converted into a continuous molar mass [*c(M)*] distribution, with accuracy dependent on the estimate for the frictional ratio (f/f_0) extracted from the experimental data.[15] The [*c(M)*] distribution gave an apparent molar mass of 31.6 kDa, showing good agreement with the monomer molar mass (31.2 kDa) calculated based on the amino-acid sequence, and further confirming that the main peak corresponds to monomeric DHDPS-A204R.

The absence of dimer was surprising, as dimeric DHDPS-A204R had been observed during the gel-filtration purification step (Figure 4.13, section 4.4.3). Therefore, a comparative experiment was performed at room temperature with an analytical gel-filtration column (Superdex 200 10/300), pre-equilibrated with 20 mM Tris.HCl pH 8.0, and where protein was diluted to ~0.01 $mg.mL^{-1}$ over the course of the chromatography. Measuring absorbance at 205 nm produced a chromatograph revealing two symmetrical elution peaks, with maxima corresponding to V_e of 13.89 and 18.82 mL **(Figure 4.9A)**. The absence of a significant peak at 10.71 mL, as was observed for tetrameric wild-type DHDPS (in chapter two, section 2.5.3), indicates the native quaternary structure is completely disrupted in mutant A204R *M. tuberculosis* DHDPS.

The chromatograph also suggest an absence of dimer, as there is no peak corresponding the molar mass (62.5 kDa) for dimeric DHDPS-A204R, which would elute at 12.05 mL based on the calibration plot **(Figure 4.9B)**. The peaks observed were converted into molar masses of 21.9 and 1.3 kDa, which unfortunately fall outside the range of the calibration standards and thus contain a large amount of uncertainty, even more than the 10 % intrinsic in this method

(discussed in chapter two, section 2.5.1). The first peak is definitely closer to molar mass predicted for the monomer (31.2 kDa) than the dimer, with -30 % deviation from the monomeric expected molar mass and therefore is consistent with the quaternary structure of DHDPS-A204R determined by AUC. The discrepancy between this result and that observed during purification is discussed in section 4.4.3. The absence of peaks corresponding to tetramer or dimer in gel-filtration chromatography indicates that secondary peak observed in AUC is most likely an artefact.

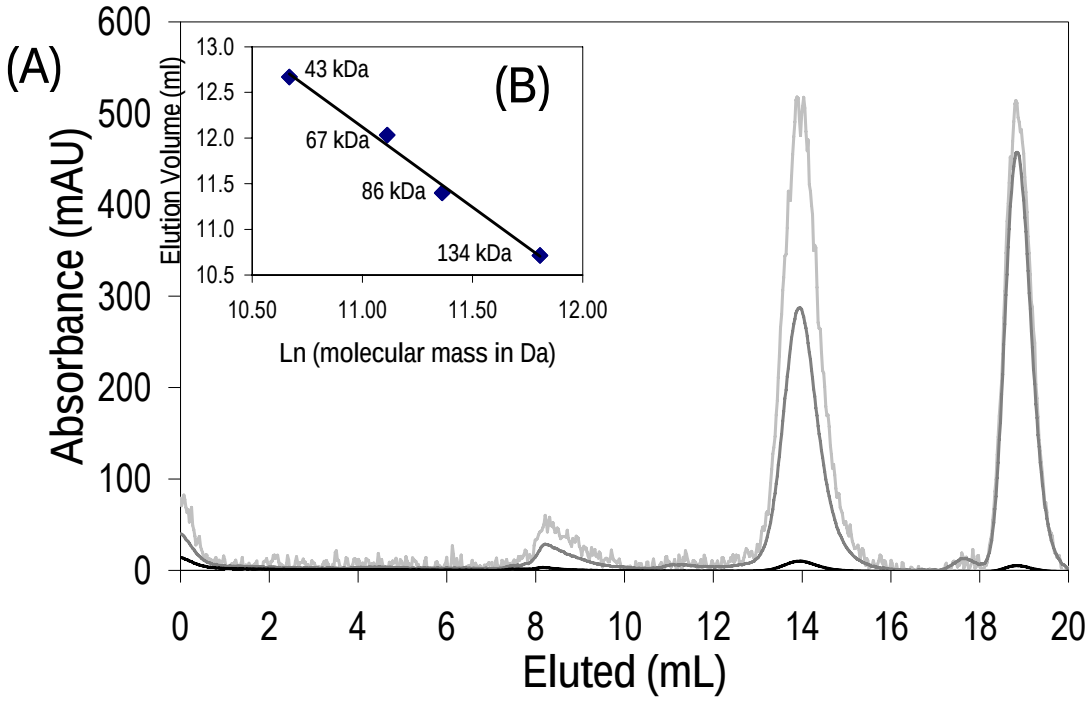


Figure 4.9: **(A) Analytical gel-filtration chromatograph of A204R *M. tuberculosis* DHDPS. The experiment was performed at room temperature in 20 mM Tris.HCl, pH 8.0, and gave elution peaks with maxima at 13.89 mL and 18.82 mL as measured by absorbance at 205 nm (light grey). Chromatographs were also recorded at 215 nm (dark grey) and 280 nm (black). (B) BSA and ovalbumin calibration standards (◆) were fitted to a linear equation.**

The peak eluting at 18.82 mL seems to correspond to something very small, possibly the cleaved His_6-tag, which has molar mass of 3.0 kDa. Proteins absorb in the 205 to 220 nm range due to their peptide bonds, and typically, absorbance measured at 215 nm is half that measured at 205 nm,[4] as is seen for the first peak with V_e of 13.89 mL. In contrast, the second

peak at 18.82 mL does not follow this pattern but rather absorbs an equivalent amount at 215 nm **(Figure 4.9A)**. Histidine has an absorbance maximum of 211 nm,[16] midway between 205 and 215 nm; thus the pattern of absorbance at the different wavelengths in the second peak is consistent with high histidine content.

Data were also collected at a higher concentration of 1.1 mg.mL^{-1} DHDPS-A204R, and fitted using a [*c(s)*] distribution, which showed two main peaks of 2.8 and 3.5 S (40 and 55 % of signal, respectively) **(Figure 4.10B)**. The fit was less than ideal, as indicated by the higher Z-test value of 20.19; however, the randomly distributed residuals and low r.m.s.d. indicate the relative goodness of the fit **(Figure 4.10A)**. Since the frictional ratio (f/f_0) reflects an estimated average of all species present and these data contain multiple peaks, the [*c(s)*] distribution cannot be converted into an accurate [*c(M)*] distribution.[17] However, the predictions by HYDROPRO suggest two oligomeric states apparently present in similar amounts at this higher concentration, suggesting the s^* values of 2.8 and 3.5 S correspond to monomer and dimer, respectively.

The first 10 radial scans were analyzed separately and gave a similar shape, indicating that the monomer-dimer equilibrium was present from the beginning of the experimental run. Subsequently, various sections of data covering different time periods were examined, and it was found that while the two peaks were not always as distinct, neither disappeared at any point. The limitations in fitting multiple peaks, especially those which overlap, meant that they merged into one broad peak in Monte-Carlo distributions* (and when confidence levels were lowered to 0.68). The concentration dependence of the monomer-dimer equilibrium is indicated by the conspicuous absence of a peak corresponding to dimer in both in the low concentration SV experiment and gel-filtration chromatography. Sedimentation equilibrium (SE) experiments were needed to provide detail about the nature of the equilibrium between oligomeric states, and confirm the absence or presence of trace amounts of tetramer.

* Monte-Carlo analysis is used to distinguish noise from true peaks. Experimental data is used to generate new data sets based on a certain confidence level, thus true peaks sharpen and noise disappears. This sharpening of true peaks was observed with the other SV data sets in this work; however, in this case because the peaks are of similar size and overlapping they merge into one broad peak.

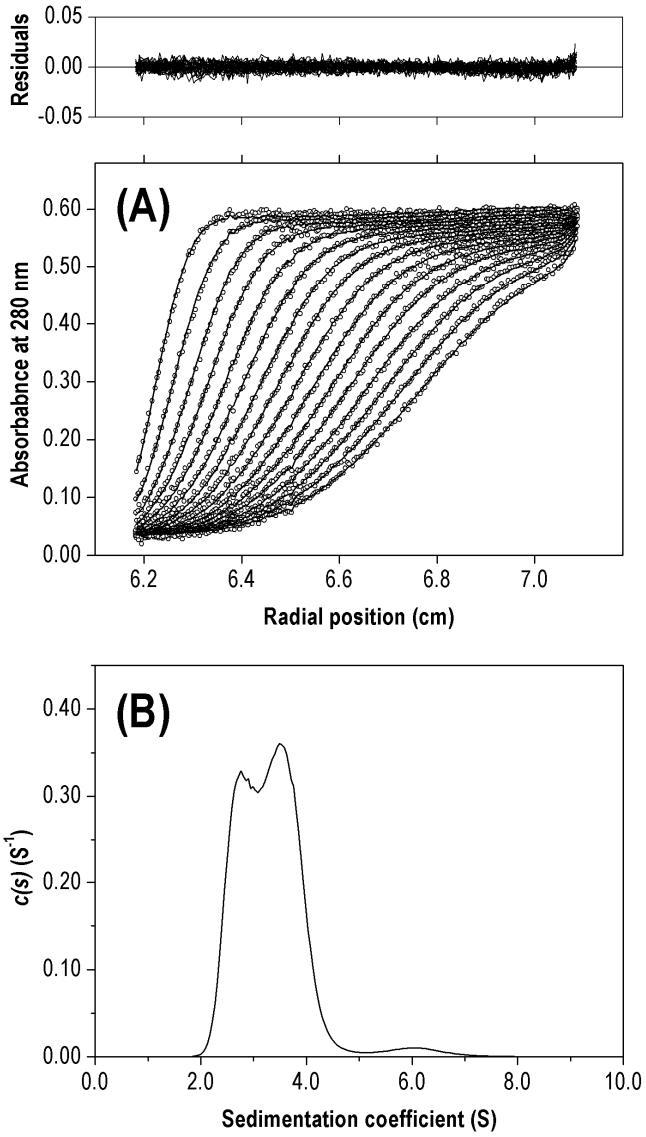


Figure 4.10: **SV analysis of DHDPS-A204R (1.1 mg.mL^{-1}) in 20 mM Tris.HCl, 150 mM NaCl, pH 8.0 at 20 ºC. (A) Absorbance at 280 nm, presented as open symbols (○), plotted at time intervals of 16 minutes is overlaid with the nonlinear least-squares fit (solid line) to a [*c(s)*] model.[12] (B) The [*c(s)*] distribution was fitted using a resolution of 200 species with $\bar{v}$ = 0.7400 mL.g^{-1}, ρ = 1.005 g.mL^{-1}, η = 1.021 cP and f/f_0 = 1.25594 and TI noise was removed. The r.m.s.d. and Z-test for the fit were 0.0040 and 20.19, respectively. The residuals are plotted as a function of radial position (cm) from the axis of rotation.**

4.4.2 Quantifying the equilibrium

To quantify the equilibrium between the oligomeric states of DHDPS-A204R, sedimentation equilibrium (SE) experiments were performed at three different rotor speeds (10000, 16000

and 23000 rpm) at three different enzyme concentrations (0.1, 0.3 and 0.9 mg.mL^{-1}). At the lower centrifugal forces used for SE experiments, diffusion starts to oppose sedimentation, and thus centrifugation for (8-16 hours) resulted in an exponential-shaped concentration distribution that was invariant with time. The unchanging concentration distribution indicated that equilibrium had been reached; that is, outward sedimentation exactly balanced inward diffusion.[18] Radial absorbance scans were recorded at 280 nm and data were analyzed at each concentration using the program SEDPHAT,[19] in similar manner as the wild-type enzyme (chapter two, section 2.5.2). Equivalent molar mass values of 37, 40 and 46 kDa were observed for the 0.1, 0.3 and 0.9 mg.mL^{-1} data sets, respectively, when fitted to a single species model. This concentration dependent increase in molar mass is characteristic of a self-associating system,[17] and suggests a shift in oligomeric species over this concentration range, as is consistent with the results from the SV experiments **(Figure 4.8** and **4.10)**.

Global analyses of all data sets were performed with the mass fixed to the calculated monomer, dimer, trimer or tetramer mass of DHDPS-A204R giving badly distributed residuals and global reduced χ^2 values of 0.57, 0.61, 5.4 and 39, respectively. None of these single species models gave a very good fit; however, of these, the best fits were obtained with the monomeric and dimeric masses.

Subsequently, the data were globally fitted to multiple self-association models; monomer-dimer, monomer-trimer, monomer-tetramer and monomer-dimer-tetramer equilibria. When the molar mass was constrained to 31 kDa, the best fit was obtained with the monomer-dimer equilibrium, yielding a dimerization dissociation constant ($K_D^{2\rightarrow1}$) of 52 μM with a global reduced χ^2 value of 0.60. Both the monomer-trimer and monomer-tetramer model gave badly distributed residuals and nonsensical dissociation constants. The monomer-dimer-tetramer model yielded a similar $K_D^{2\rightarrow1}$ of 51 μM, with global reduced χ^2 value of 0.60; however, the tetramerization dissociation constant ($K_D^{4\rightarrow2}$) was in the molar concentration range, several 1000-fold higher than the highest loading concentration of 0.9 mg.mL^{-1} (28.8 μM). This indicated that no tetramer is apparent in the concentration range examined, and that the appropriate model for DHDPS-A204R over the concentrations investigated is the monomer-dimer equilibrium. The fit was optimized for the monomer-dimer model, yielding a dimerization dissociation constant ($K_D^{2\rightarrow1}$) of 60 μM with a χ^2 value of 0.26. When the molar mass was floated, a fit with a χ^2 value of 0.25 was obtained for the monomer-dimer model with buoyant molar mass for the monomer of 29.7 kDa and $K_D^{2\rightarrow1}$ of 51 μM **(Figure 4.11)**.

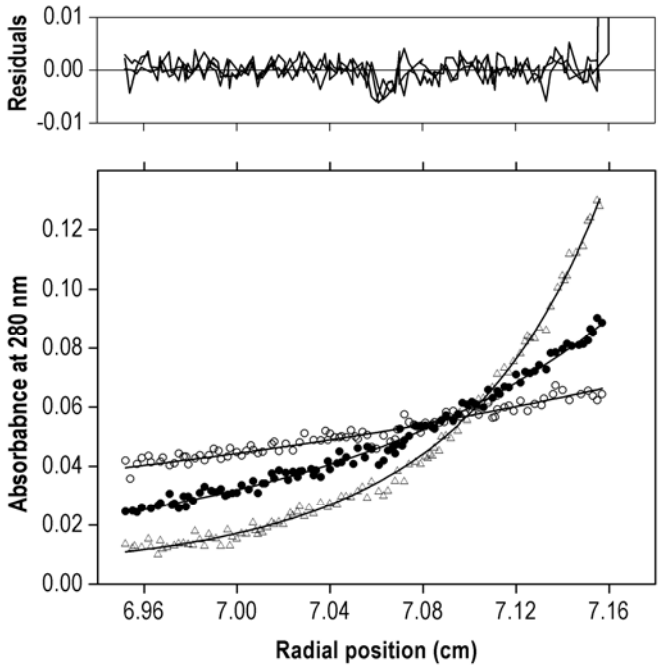


Figure 4.11: **DHDPS-A204R sedimentation equilibrium data recorded at 280 nm and 20 ºC, in 20 mM Tris.HCl, 150 mM NaCl, pH 8.0. Representative data are shown for initial protein concentration of 0.1 mg.mL^{-1} after reaching equilibrium at 10000 rpm (○), 16000 rpm (●), and 23000 rpm (Δ) overlaid with the nonlinear least-squares fit (solid line) to a monomer-dimer self-associating model yielding a $K_D^{2\rightarrow1}$ of 51 µM and buoyant molar mass of 29.7 kDa. The residuals are plotted as a function of radial position (cm) from the axis of rotation.**

The $K_D^{2\rightarrow1}$ is the concentration at which the amount of monomer equals the amount of dimer, and can also be expressed as 1.6 mg.mL^{-1}. This is slightly higher than the 1.1 mg.mL^{-1} that gave similar sized peaks for monomer and dimer with the SV experimental approach. Since $K_D^{2\rightarrow1}$ describes the ratio of monomer to dimer it can be used to predict the fraction of monomer (f_m) present at any protein concentration, as described with **Equation 4.1**:

$$f_m = \frac{-K_D^{2\rightarrow1} + \sqrt{(K_D^{2\rightarrow1})^2 + 8(K_D^{2\rightarrow1}) \cdot P_T}}{4P_T} \qquad \textbf{Equation 4.1}\ ^{20}$$

Here P_T is the total protein concentration, expressed in molar concentration units,[20] thus by converting 1.1 mg.mL^{-1} to 35 µM, the fraction of monomer can be calculated as 0.56, or 56 %, with the remaining 44 % expected to be dimer. In the SV data analysis of DHDPS-A204R at 1.1 mg.mL^{-1} (section 4.4.2, Figure 4.10B) the opposite trend was observed, as the

monomer peak was smaller than the dimer peak (40 and 55 %, respectively); however, these values are still relatively close to those predicted by the $K_D^{2\rightarrow1}$ determined using the SE experimental approach. The difference probably reflects the intrinsic inaccuracy when fitting interacting systems with the Lamm equation for SV data analysis, and highlights the importance of using SE analysis to characterize rapidly equilibrating, self-associating systems.

4.4.3 Factors affecting equilibrium

During purification, only dimeric A204R *M. tuberculosis* DHDPS was observed on the preparative gel-filtration column (Superdex 200 16/60) **(Figure 4.13)**. The lower resolution preparative-type column meant yeast alcohol dehydrogenase needed to be included to generate the four point calibration curve and showed the single slightly asymmetrical peak corresponded to 57.3 kDa. This suggested a mainly dimeric DHDPS-A204R, with a small fraction of monomer. The gel-filtration purification step was performed at 4 ºC, in storage buffer containing pyruvate, and the concentration of enzyme collected was usually ~0.5 mg.mL^{-1} (section 4.2.2, Table 4.2), thus less than the $K_D^{2\rightarrow1}$ of 1.6 mg.mL^{-1} determined in 20 mM Tris.HCl, 150 mM NaCl, pH 8.0 at 20 ºC. Using equation 4.1 (section 4.4.2), the $K_D^{2\rightarrow1}$ predicts that 70 % of DHDPS-A204R will adopt a monomeric quaternary structure at 0.5 mg.mL^{-1}, therefore, the relative absence of monomer suggested that the monomer-dimer equilibrium of DHDPS-A204R has been shifted towards dimer by one or more of the storage buffer components. The lower temperature of 4 ºC maintained during purification may have also influenced subunit association, as temperature effects have been observed for several oligomeric enzymes.[21,22]

To investigate the effect of the buffer components on equilibrium, SV experiments were performed at 20 ºC with 0.6 mg.mL^{-1} DHDPS-A204R in storage buffer [20 mM Tris.HCl, 250 mM NaCl, 5 % glycerol, 2 mM β-mercaptoethanol (β-ME), 10 mM pyruvate, pH 8.0]. Due to the complexity of the buffer in this case, the computational processing of data relied on estimates for protein and solution properties. The limitations of the SEDNTERP program meant contributions to buffer density or viscosity from the low concentrations of pyruvate and β-ME had to be ignored. Solvent density (ρ) and viscosity (η) were calculated as 1.023 g.mL^{-1} and 1.193 cP, respectively, for a pH 8.0 buffer composed of 20 mM Tris.HCl, 250 mM NaCl and 5 % (0.68 M) glycerol. The partial specific volume ($\bar{v}$) of 0.7400 mL.g^{-1} used in data analysis was also an estimate. The presence of glycerol is known to impact the $\bar{v}$ of proteins,

increasing protein volume at acidic pH by preferential hydration,[23] and decreasing protein volume at neutral pH through compression of the protein interior.[24] Parallel SV experiments performed with wild-type *M. tuberculosis* DHDPS (1.1 mg.mL^{-1}) in the same storage buffer and allowed the inaccuracy introduced by these estimates to be monitored.

HYDROPRO predicted sedimentation coefficients for monomer, dimer and tetramer, in 20 mM Tris.HCl, 250 mM NaCl, 5 % glycerol to be 2.2 S, 3.5 S and 5.5 S, respectively. Using SEDFIT, a [*c(s)*] distribution was fitted to the wild-type data set and gave a peak of 5.3 S **(**see **Appendix H)**, which reflects the previously demonstrated tetrameric quaternary structure of wild-type *M. tuberculosis* DHDPS (chapter two, section 2.6.1). The closeness of this value to the sedimentation coefficient predicted by HYDROPRO indicates that the estimates for solution and protein properties have introduced minimal inaccuracy. Furthermore, when s^* was standardized to conditions corresponding to pure water at 20 °C, $s°_{20,w}$, using SEDNTERP,[13] a value of 6.8 S was obtained, which is similar to 6.5 S found for the wild-type enzyme in SV experiments performed in a different buffer.

Analysis of the data set for 0.6 mg.mL^{-1} DHDPS-A204R in storage buffer revealed a single peak of 3.2 S **(Figure 4.12B)**, which is close to the 3.5 S predicted for a dimer by HYDROPRO. The random distribution of the residuals and low r.m.s.d. of 0.0036 indicates the quality of the fit **(Figure 4.12A)**. The $s°_{20,w}$ of 4.1 S calculated from s^* is significantly different from 2.8 S obtained in the other buffer [20 mM Tris.HCl, 150 mM NaCl, pH 8.0], which corresponded to monomer. The [*c(M)*] distribution derived from the [*c(s)*] distribution using SEDFIT gave an apparent molar mass of 62.4 kDa, which is consistent with the predicted molar mass for a dimer of 62.5 kDa. These three lines of evidence show that the single peak observed at 0.6 mg.mL^{-1} in storage buffer corresponds to a dimer, **(Figure 4.12B)** and the absence of any distinguishable monomer peak confirms that one or more components of the storage buffer has shifted the monomer-dimer equilibrium of DHDPS-A204R towards dimer.

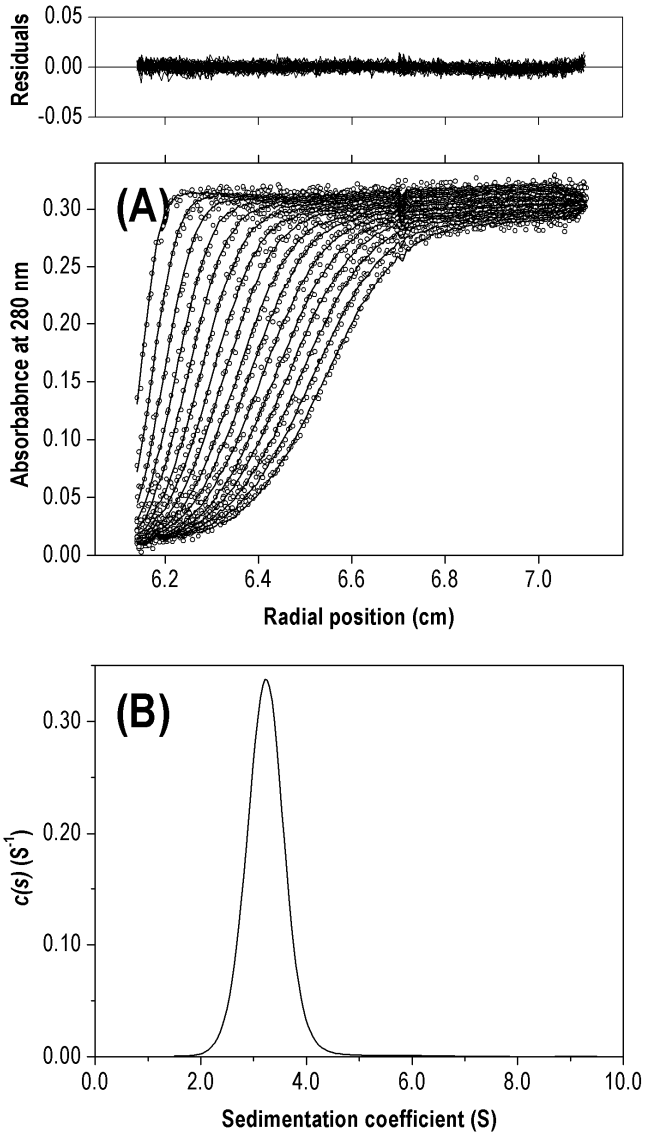


Figure 4.12: **Sedimentation velocity analysis of DHDPS-A204R (0.6 mg.mL^{-1}) in storage buffer [20 mM Tris.HCl, pH 8.0, 250 mM NaCl, 5 % glycerol, 2 mM β-ME, and 10 mM pyruvate] at 20 ºC. (A) Absorbance at 280 nm, presented as open symbols (○), plotted at time intervals of 12 minutes is overlaid with the nonlinear least-squares fit (solid line) to a [*c(s)*] model.[12] (B) The fit was obtained using a resolution of 200 species with $\bar{\nu}$ = 0.7400 mL.g^{-1}, ρ = 1.023 g.mL^{-1}, η = 1.193 cP and f/f_0 = 1.20145, and TI noise was removed. The r.m.s.d. and Z-test for the fit were 0.0036 and 12.44, respectively. The residuals are plotted as a function of radial position (cm) from the axis of rotation.**

Two components in the storage buffer were identified as possibly causing the oligomeric change. The binding to the enzyme of substrate has been reported in the literature to induce

dimerization.[25,26] The presence of anti-chaotropic agents, such as glycerol have also been observed to affect oligomeric equilibrium.[27,28] To investigate which storage buffer component was responsible for the observed shift in oligomeric structure, a gel-filtration run was performed in storage buffer lacking pyruvate, re-calibrated with BSA, ovalbumin and alcohol dehydrogenase **(Figure 4.13B)**. The elution peaks of all proteins shifted in the absence of pyruvate, however the peak of DHDPS-A204R shifted more significantly **(**black line, **Figure 4.13A)**, giving an elution volume of 95.34 mL.

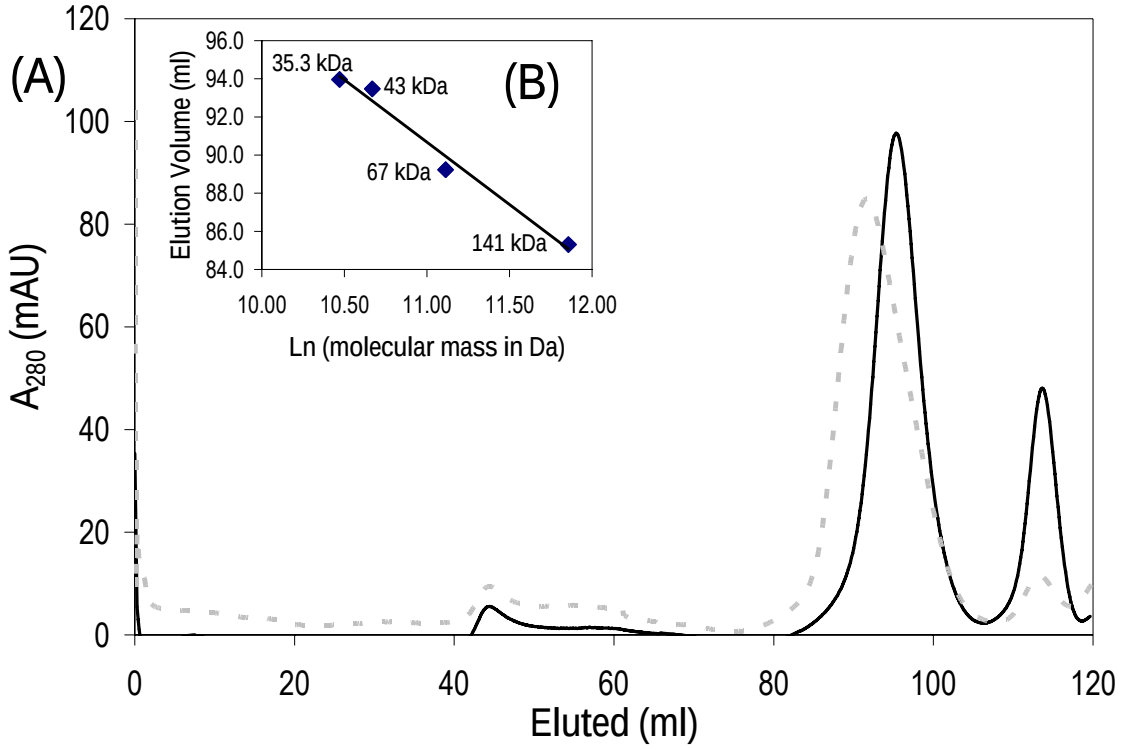


Figure 4.13: (A) Gel-filtration chromatography of A204R *M. tuberculosis* DHDPS, performed at 4 ºC in a pH 8.0 buffer [20 mM Tris.HCl, 250 mM NaCl, 5 % glycerol, 2 mM β-ME] with 10 mM pyruvate (light grey dashes) and without pyruvate (black) giving elution peaks corresponding to dimer and monomer, respectively. The second, less significant, peak is likely the cleaved His_6-tag, which gave a small peak in most purification gel-filtration runs. (B) Calibration standards of BSA, ovalbumin, and alcohol dehydrogenase (◆) in pH 8.0 buffer without pyruvate.

The slightly asymmetrical peak **(Figure 4.13A)** showing DHDPS activity corresponded to a molar mass of 29.3 kDa, as determined with the new calibration plot **(Figure 4.13B)** and

indicated that in the absence of pyruvate, DHDPS-A204R is mostly monomeric. Since the preparative-type column has a lower resolution, this profile may fit the prediction of 70 % monomeric DHDPS-A204R at 0.5 mg.mL^{-1} made using the $K_D^{2\rightarrow1}$ determined by SE analysis in 20 mM Tris.HCl, 150 mM NaCl, pH 8.0 at 20 ºC.

Further experiments were performed at 4 °C with an analytical gel-filtration column (Superdex 200 10/300) to explore the effect of pyruvate concentration on the oligomeric structure of DHDPS-A204R. Relatively low concentration samples (0.16 mg.mL^{-1}) were loaded (250 µL) onto the column which had been pre-equilibrated with buffer containing varying concentrations of pyruvate **(Figure 4.14)**. At these low concentrations the dimeric and monomeric peaks could be distinguished and thus their V_e could be converted into molecular masses using BSA and ovalbumin calibration standards **(Table 4.4)**. The *o*-aminobenzaldehyde assay showed DHDPS activity in both monomer and dimer peaks. The monomer-dimer equilibrium shifts from majority monomer to majority dimer at low pyruvate concentrations, but is only slightly affected by further increasing pyruvate concentration, consistent with stoichiometric binding of pyruvate stabilising the dimeric state.

Table 4.4: **The molecular masses and corresponding quaternary structures of A204R *M. tuberculosis* DHDPS in pH 8.0 buffer (20 mM Tris.HCl, 150 mM NaCl) and varying pyruvate concentrations. (Bold denotes majority species).**

[Pyruvate] mM	Molar mass (kDa)		Quaternary structure	
0.0	57	34	dimer	monomer
0.5	58	33	dimer	monomer
1.0	56	32	dimer	monomer
5.0	56	31	dimer	monomer
10.0	62	36	dimer	monomer

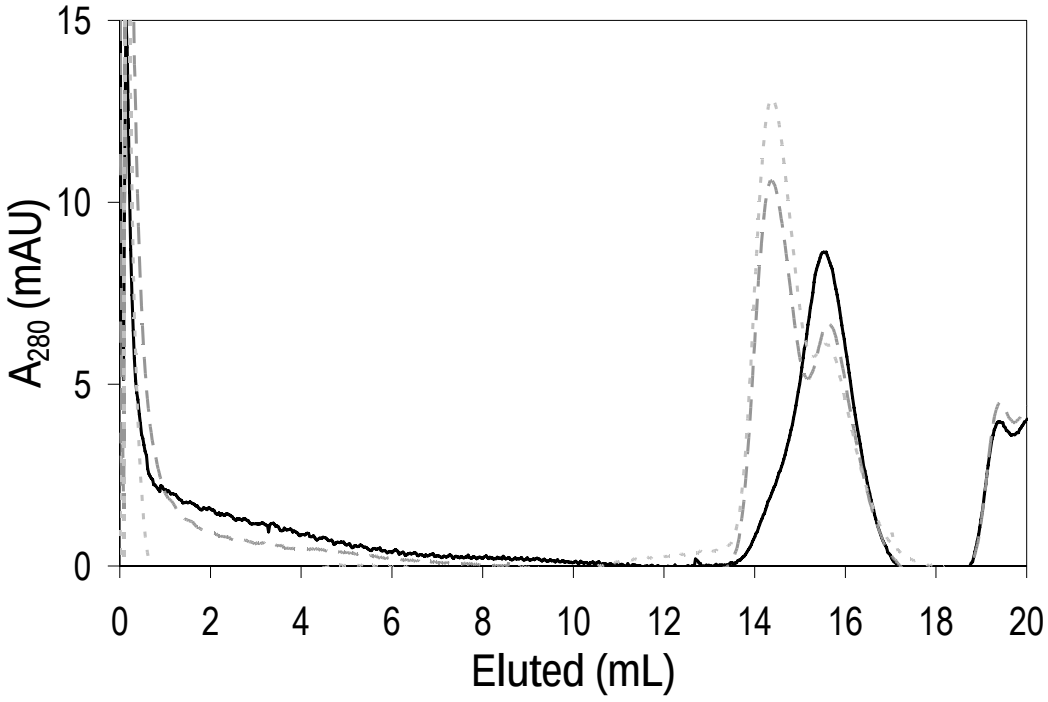


Figure 4.14: The effect of pyruvate concentration on A204R *M. tuberculosis* DHDPS shown by analytical gel filtration performed at 4 ºC with 10 mM pyruvate (light grey dots), 0.5 mM pyruvate (dark grey dashed) and without pyruvate (black) in a pH 8.0 buffer [20 mM Tris.HCl, 150 mM NaCl]. The column was loaded with 0.04 mg of A204R DHDPS (250 µL of 0.16 mg.mL^{-1}), thus eluted at a concentration of ~0.01 mg.mL^{-1} (0.04 mgs over ~4 mL).

4.5 The crystal structure of A204R *M. tuberculosis* DHDPS

X-ray crystallography provides the level of detail required for the similarities and differences between the interface mutant DHDPS-A204R and wild-type *M. tuberculosis* DHDPS to be understood in the context of disrupted quaternary structure

4.5.1 Crystallization, diffraction data collection and processing

Crystallization experiments were undertaken by our collaborators at the EMBL Hamburg Outstation using protein purified during the course of this thesis. The first crystals were observed in the condition of 2.0 M $(NH_4)_2SO_4$, and 100 mM Tris.HCl, pH 8.5, at room temperature, and subsequently, this condition was optimized to 2.0 M $(NH_4)_2SO_4$ and 100 mM sodium acetate pH 5.5.

X-ray diffraction data were collected by our collaborators on the I911-3 beamline at the MAX-lab synchrotron in Lund (Sweden) using a MARMosaic CCD detector. The data were indexed and integrated using DENZO and scaled with SCALEPACK.[29] Data collection and processing parameters are given in **Table 4.5**, and further details on data collection are given in chapter six, section 6.10.1.

Table 4.5: Collection and processing statistics for X-ray crystallography of DHDPS-A204R. Values in parentheses are for the highest resolution bin.

No. of crystals	1
Wavelength (Å)	1.000
Crystal-detector distance (mm)	225
Rotation range per image (°)	0.5
Exposure time per image	30 sec
Total rotation range (°) [a]	110.5
Resolution range (Å)	99.00-2.00 (2.03-2.00)
Space group	$P4_12_12$
Unit cell parameters (Å, °)	188.83, 188.83, 130.43, 90.0, 90.0, 90.0
Mosaicity (°)	0.26
Total no. of observations	1,458,570
Total no. of reflections	747,046
Unique reflections	157,209
Rejected reflections	9,630
Redundancy	4.8 (4.3)
I/σ (I)	14.1 (2.5)
Completeness (%)	99.2 (97.6)
R_{merge} (%) [b]	11.8 (58.9)
Overall B-factor from Wilson plot (Å^2)	23.1
Optical resolution (Å)	1.54

[a] Only the first 60° of data were used for final data processing due to radiation damage.

[b] $$R_{merge} = \frac{\sum_{hkl}\sum_{i}\left|I_i(hkl) - \langle I(hkl)\rangle\right|}{\sum_{hkl}\sum_{i} I_i(hkl)}$$

4.5.2 Structural determination and initial refinement

Our collaborators solved the structure by molecular replacement using the program MOLREP,[30] and a dimer consisting of the chains *a* and *b* of wild-type *M. tuberculosis* DHDPS (PDB entry 1XXX) as a search model. The crystal structure of DHDPS-A204R was solved in the space group $P4_12_12$ and contained three independent dimers **(Figure 4.15A)**, with no tetramers seen in crystal packing, providing unequivocal evidence that the tetrameric structure is disrupted by the alanine to arginine mutation. After correctly orienting and positioning the three dimers, the *R*-factor was already as low as 36.8 %. One round of rigid body and restrained refinement, using the program REFMAC5,[30,31] reduced the *R*- and free *R*-factor to 32.0 % and 27.8 %, respectively. This partially refined data was then sent to our lab for further refinement, which I carried out with support from Prof. Geoff Jameson (University of Massey) and Dr Sean Devenish.

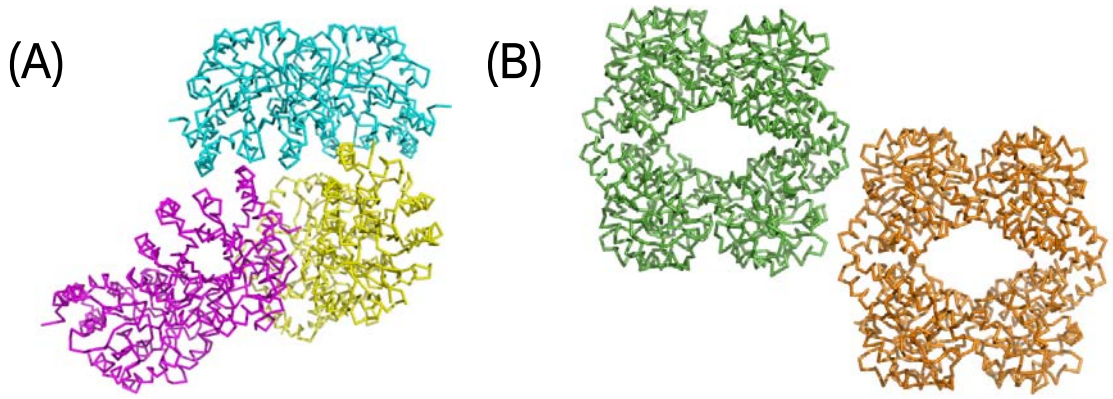


Figure 4.15: **Comparison of the crystal packing of (A) mutant A204R and (B) wild-type *M. tuberculosis* DHDPS (PDB entry 1XXX) in the unit cell, showing the quaternary structures of dimer and tetramer, respectively.**

4.5.3 Further structural refinement

Because of the multiplicity in the asymmetric unit, non-crystallographic similarity (NCS) constraints could be used with rigid-body refinement. After two rounds of rigid body refinement with NCS constraints, the *R*- and free *R*-factors were 26.3 % and 24.4 %, respectively. Anisotropic displacement parameters were incorporated into refinement using the TLS (translation, rotation, screw-rotation) model with REFMAC5, which can be visualized as six thermal ellipsoids, each constraining a monomer.[30,32] Manual model corrections were implemented using the program WINCOOT,[33] and new models were further

refined using REFMAC5. During the manual model corrections, the alanine of the wild-type enzyme was replaced by an arginine residue at the position of the mutation. The mutation was clearly visible in the electron density in four of the six monomers and partially visible (due to disorder in the side chain) in two monomers.

During refinement, well-defined electron density was observed near the ε-amino group of the active-site lysine residue, K171, in all six monomers. The size, shape and orientation of this density was consistent with a pyruvate-adduct **(Figure 4.16A)**, as had been observed in structures solved for DHDPS from other species, such as the recently determined structure of *E. coli* DHDPS.[34] Consequently, pyruvate was modelled into the final structure as a Schiff base covalently bound to the ε-amino group of K171 **(Figure 4.16B)**. That pyruvate would be bound in the active site was not unexpected given that the enzyme had been purified and stored in a buffer containing 10 mM pyruvate, which was also observed to result in formation of a pyruvate-adduct in *C. glutamicum* DHDPS.[6]

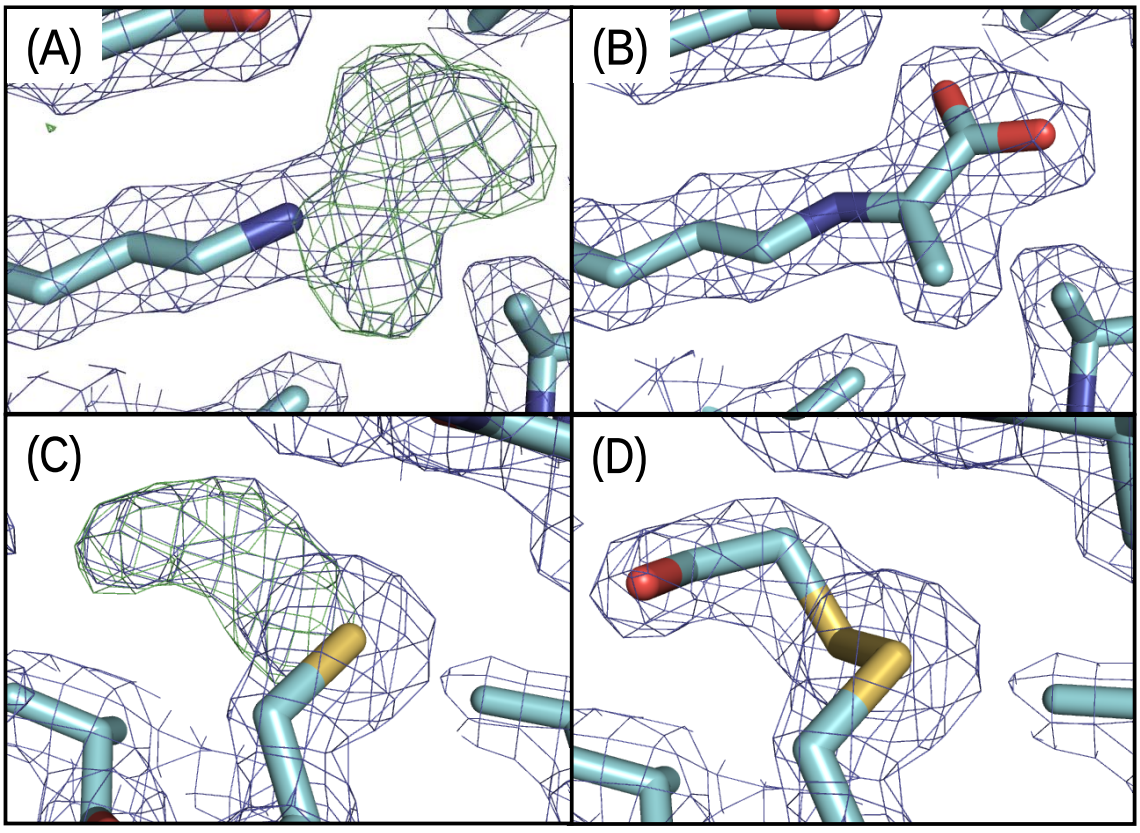


Figure 4.16 **Omit maps of the density seen at (A) K171 and (C) C248, shown for chain A, analogous to that observed in other monomers, and the corresponding fit after modelling adducts with (B) pyruvate and (D) β-mercaptoethanol (β-ME).**

Additional density was also noticed near the surface cysteine residue, C248, in all monomers **(Figure 4.16C)**. In the wild-type *M. tuberculosis* DHDPS (PDB 1XXX) DTT molecules were found to be covalently bound to C248,[35] but DTT did not fit the density observed in the crystal structure of DHDPS-A204R. Subsequently, another reducing agent used in purification, β-mercaptoethanol (β-ME), was modelled into the structure and found to fit the density well **(Figure 4.16D)**.

The final refinement rounds involved placement of solvent molecules using WINCOOT, followed again by manual corrections and further refinement with REFMAC5. Structural quality was assessed using SFCHECK,[36] and the structural validation tools of WINCOOT. In some cases, the side chains of amino acids could not be clearly observed due to insufficient electron density, so the occupancy of these atoms was lowered. The maximum r.m.s.d. for the subunits (all atoms) in the asymmetric unit was 0.38 Å, as calculated by SUPERPOSE.[30] Final refinement statistics are presented in **Table 4.6**.

Table 4.6: Refinement statistics for the DHDPS-A204R crystal structure. Values in parentheses are for the highest resolution bin.

Resolution (Å)	2.1
Resolution range (Å)	28.51-2.10 (2.213-2.10)
R_{free} (outer shell) [a]	0.204 (0.234)
R_{cryst} (outer shell) [b]	0.178 (0.196)
Residues/water molecules	1766/1914
r.m.s.d. from ideal geometry	
Bond lengths (Å)	0.015
Bond angles (°)	1.52

[a] R_{free} based on 5 % of the total reflections excluded from the refinement.

[b] $R_{cryst} = \frac{\sum\left(\left|F_{obs}\right| - \left|F_{calc}\right|\right)}{\sum\left|F_{obs}\right|}$

4.5.4 General features of the structure

The final model contained six monomers in the asymmetric unit, with 1766 amino-acid residues in total, 1914 water molecules, six β-mercaptoethanol (β-ME) moieties covalently bound to C248, six pyruvate moieties covalently bound to K171, along with one non-

covalently bound β-ME and pyruvate from the purification buffer, and 16 sulfate molecules, 24 glycerol molecules, eleven acetate molecules and three chloride ions from the crystallization buffer. The non-covalently bound pyruvate and β-ME were both on the outer surface of the enzyme, distal to the active site, and unlikely to have functional significance. The model includes all amino-acid residues, with the exception of the first 8 to 10 amino acids at the N-terminus of each chain (including the G-A-M-A remaining after cleavage of the His_6-tag), which are not visible in the electron density.

Almost all non-glycine residues, 99.8 %, fell within the most favoured or additionally allowed regions of the Ramachandran plot.[37] The only offending residue was the interdigitating tyrosine, Y117, which reaches from one monomer to complete the adjacent active site. This disfavoured orientation has been associated with the tyrosine residue's catalytic function and has been observed in wild-type *M. tuberculosis* DHDPS,[35] and in DHDPS from other species.[6,25,38-45] In this structure, three of the six Y117 residues were in the disallowed region of the plot, while the others were found just inside the additionally allowed region, due to slight adjustments in orientation. This may reflect either the tightness of restraints being used during refinement, or slight variation between the monomers in the asymmetric unit.

4.5.5 Alignment of the DHDPS-A204R and the wild-type structures

Superpositioning of all atoms in monomer *a* with the wild-type *M. tuberculosis* DHDPS structure (1XXX), and the structure of wild-type soaked with pyruvate, gave r.m.s.d. values of 0.288 Å and 0.289 Å, respectively. These values are comparable to those of alignments of the monomers of DHDPS-A204R within the asymmetric unit (section 4.5.3), also determined using SUPERPOSE. This high degree of structural similarity indicates that the mutation of alanine to arginine has not introduced any gross disruptions to the secondary or tertiary structure of the enzyme.

An alignment of the dimer composed of chains *a* and *b* with the corresponding chains in 1XXX also yielded a low r.m.s.d. of 0.378 Å for all atoms, which is perhaps unsurprising, considering chains *a* and *b* of 1XXX were used as the search model for the DHDPS-A204R structure, and resulted in a low initial *R*-factor after a single round of rigid body refinement. An analogous alignment using the wild-type structure with pyruvate bound at the active site gave a slightly better fit, with an r.m.s.d. of 0.337 Å **(Figure 4.17A)**, perhaps reflecting the

subtle shift noted in active-site residues, as discussed in chapter two, section 2.6.3. A closer look at the active site **(Figure 4.17B)**, shows little change in the substrate binding and catalytic residues and will be discussed in detail in the next section.

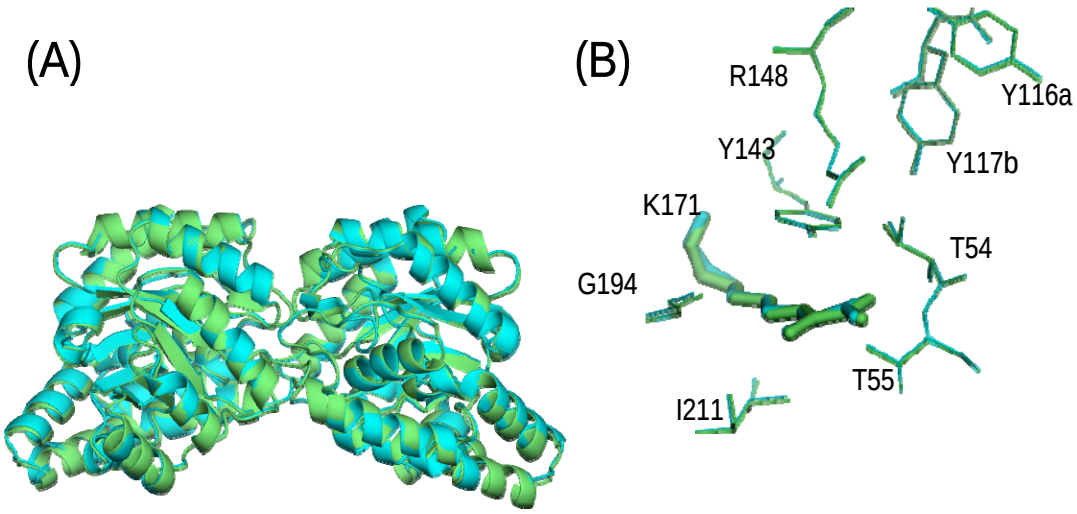


Figure 4.17: (A) Alignment of the dimer A204R *M. tuberculosis* DHDPS (in blue) with two subunits from wild-type enzyme (in green), both with pyruvate covalently bound to K171. (B) Overlay of the active site showing the close alignment of key residues, which was produced from an alignment of single monomers of the structures.

4.5.6 The active site

The active sites of wild-type tetrameric *M. tuberculosis* DHDPS and dimeric DHDPS-A204R are remarkably similar **(Figure 4.17B)**, consistent with their similar kinetically quantified catalytic activity. The disrupted quaternary structure mutants of *E. coli* DHDPS all had greatly reduced activity, which was attributed in part to the presence of covalent adduct, α-ketoglutarate, in the active site.[5,46] The presence of pyruvate rather than α-ketoglutarate at the active-site lysine residue of DHDPS-A204R is consistent with the interface mutant having comparable activity to the wild-type enzyme.

The proton-relay network, composed of two tyrosines (Y143, Y117) and a threonine (T54), as noted in chapter two, appears to be disrupted in the crystal structure of wild-type *M. tuberculosis* DHDPS (PDB entry 1XXX). However, in soaking experiments with pyruvate and wild-type *M. tuberculosis* DHDPS, the hydrogen bond between Y143-OH and T54-OH appears to be re-established, thus restoring the proton-relay network. Analogous to the

pyruvate-bound structure of *M. tuberculosis* DHDPS, the distances between Y143-OH and T54-OH indicates hydrogen bond formation in the pyruvate-bound DHDPS-A204R crystal structure **(**representative numbers from monomer *a* shown in **Table 4.7)**. The atomic distances of the proton-relay network determined for DHDPS-A204R were similar in all six monomers, indicating that the proton-relay motif remains intact.

Table 4.7 Comparison of key atomic distances within the proton-relay network from monomer *a* of crystal structures solved for wild-type and A204R *M. tuberculosis* DHDPS and wild-type *E. coli* DHDPS containing a pyruvate adduct.

Atomic distances for :		*M. tuberculosis* DHDPS[a]	*M. tuberculosis* DHDPS-A204R	Atomic distances for:		*E. coli* DHDPS[b]
Y143-OH	K171-Nϵ[c]	3.3 Å	3.6 Å	Y133-OH	K161-Nϵ[c]	3.5 Å
Y143-OH	T54-OH	3.0 Å	3.2 Å	Y133-OH	T44-OH	2.8 Å
T54-OH	Y117-OH	2.6 Å	2.6 Å	T44-OH	Y107-OH	2.6 Å

[a] Structure of *M. tuberculosis* DHDPS with pyruvate bound was solved by our collaborators and has not yet been deposited in the PDB.

[b] Structure of *E. coli* DHDPS with pyruvate bound is PDB entry 3DU0.

[c] The ϵ-amino group of the key lysine residue that forms a Schiff base with pyruvate.

For DHDPS-A204R, in contrast to the wild-type structure with a pyruvate adduct, the distance between Y143-OH and the ϵ-amino group of K171 increased from 3.3 to 3.6 Å **(Table 4.7)**. This difference of 0.3 Å is near the limit of what can be consider significant with this data set. It results from slight re-orientations of K171 and Y143 **(Figure 4.17B)**. Interestingly, the shifting of the key active-site lysine away from the proton-relay network upon pyruvate binding has also been observed in wild-type tetrameric *E. coli* DHDPS structures (PDB code 1YXC,[44] 3DU0,[34]). Thus the re-positioning of the proton-relay network may be related to the binding of pyruvate. Regardless, the re-orientations of Y143 and K171 in this crystal structure is slight and unlikely to be due to the mutation, distal to the active site. Thus, the differences observed in the kinetic parameters of DHDPS-A204R and wild-type *M. tuberculosis* DHDPS can be attributed to disrupted quaternary structure rather than changes at the active site propagated by the mutation of alanine to arginine.

4.6 Summary

A comprehensive biophysical, biochemical and structural characterization of mutant A204R *M. tuberculosis* DHDPS revealed an enzyme with comparable activity to wild-type tetrameric DHDPS, lowered thermal stability, and a completely disrupted quaternary structure. Using AUC and gel-filtration chromatography, this interface mutant was shown to exist in a monomer-dimer equilibrium, affected by enzyme concentration and by the presence of the first substrate, pyruvate. The crystal structure of DHDPS-A204R, solved with pyruvate bound in the active site, provided further confirmation of the disrupted quaternary structure.

Additionally, the crystal structure showed DHDPS-A204R to have tertiary and secondary structures identical to the wild-type enzyme, and the active site was revealed to be intact in the static structure. This meant that DHDPS-A204R provided a valid model for determining the role of quaternary structure in *M. tuberculosis* DHDPS, uncomplicated by changes to secondary and tertiary structure.

In summary, the dimeric *M. tuberculosis* DHDPS was successfully created and characterized. Contrary to the hypothesis outlined in chapter one, and unlike its *E. coli* counterpart, the dimeric enzyme retained catalytic competence. This unequivocally rules out interface disruption as a mode of inhibition for *M. tuberculosis* DHDPS. It also demands a re-evaluation of the role of quaternary structure in DHDPS function, which will be discussed in the concluding chapter.

4.7 References

1. Ericsson, U. B., Hallberg, B. M., Detitta, G. T., Dekker, N. & Nordlund, P. Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Analytical biochemistry* **357**, 289-98 (2006).

2. Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein expression and purification* **41**, 207-34 (2005).

3. Good, N. E., Winget, G. D., Winter, W., Connolly, T. N., Izawa, S. & Singh, R. M. Hydrogen ion buffers for biological research. *Biochemistry* **5**, 467-77 (1966).

4. Scopes, R. K. *Protein purification: principles and practice* (Springer-Verlag, New York, 1987).

5. Griffin, M. D., Dobson, R. C., Pearce, F. G., Antonio, L., Whitten, A. E., Liew, C. K., Mackay, J. P., Trewhella, J., Jameson, G. B., Perugini, M. A. & Gerrard, J. A. Evolution of quaternary structure in a homotetrameric enzyme. *Journal of molecular biology* **380**, 691-703 (2008).

6. Rice, E. A., Bannon, G. A., Glenn, K. C., Jeong, S. S., Sturman, E. J. & Rydel, T. J. Characterization and crystal structure of lysine insensitive *Corynebacterium glutamicum* dihydrodipicolinate synthase (cDHDPS) protein. *Archives of biochemistry and biophysics* **480**, 111-21 (2008).

7. Wolterink-van Loo, S., Levisson, M., Cabrieres, M. C., Franssen, M. C. & van der Oost, J. Characterization of a thermostable dihydrodipicolinate synthase from *Thermoanaerobacter tengcongensis*. *Extremophiles* **12**, 461-9 (2008).

8. Dobson, R. C., Gerrard, J. A. & Pearce, F. G. Dihydrodipicolinate synthase is not inhibited by its substrate, (*S*)-aspartate β-semialdehyde. *The Biochemical journal* **377**, 757-62 (2004).

9. Cornish-Bowden, A. *Fundamentals of enzyme kinetics* (Butterworths, London ; Boston, 1979).

10. Dixon, M. & Webb, E. C. *Enzymes* (Longman, London, 1979).

11. Cornish-Bowden, A. *Fundamentals of enzyme kinetics* (Portland Press, London, 1995).

12. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modelling. *Biophysical journal* **78**, 1606-19 (2000).

13. Laue, T. M., Shah, D. B., Ridgeway, T. M. & Pelletier, S. L. *Analytical ultracentrifugation in biochemistry and protein science* (The Royal Society of Chemistry, Cambridge, 1992).

14. Garcia De La Torre, J., Huertas, M. L. & Carrasco, B. Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophysical journal* **78**, 719-30 (2000).

15. Lebowitz, J., Lewis, M. S. & Schuck, P. Modern analytical ultracentrifugation in protein science: a tutorial review. *Protein science: a publication of the protein society* **11**, 2067-79 (2002).

16. Mooz, E. D. Amino Acids in *Practical handbook of biochemistry and molecular biology* (ed. Fasman, G. D.) 3-97 (CRC Press, Boca Raton, FL, 1989).

17. Balbo, A. & Schuck, P. Analytical ultracentrifugation in the study of protein self-association and heterogeneous protein-protein interactions in *Protein-protein interactions: a molecular cloning manual* (ed. Golemis, E.) 253-78 (Cold Spring Harbor Laboratory Press, New York, 2005).

18. Ralston, G. *Introduction to analytical ultracentrifugation* (Beckman Instruments, Fullerton, CA, 1993).

19. Vistica, J., Dam, J., Balbo, A., Yikilmaz, E., Mariuzza, R. A., Rouault, T. A. & Schuck, P. Sedimentation equilibrium analysis of protein interactions with global implicit mass conservation constraints and systematic noise decomposition. *Analytical biochemistry* **326**, 234-56 (2004).

20. Goodman, M. D. & Hargrove, M. S. Quaternary structure of rice nonsymbiotic hemoglobin. *The Journal of biological chemistry* **276**, 6834-9 (2001).

21. Jaenicke, R. Protein structure and function at low temperatures. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **326**, 535-51; discussion 51-3 (1990).

22. Traut, T. W. Dissociation of enzyme oligomers: a mechanism for allosteric regulation. *Critical reviews in biochemistry and molecular biology* **29**, 125-63 (1994).

23. Gekko, K. & Timasheff, S. N. Mechanism of protein stabilization by glycerol: preferential hydration in glycerol-water mixtures. *Biochemistry* **20**, 4667-76 (1981).

24. Priev, A., Almagor, A., Yedgar, S. & Gavish, B. Glycerol decreases the volume and compressibility of protein interior. *Biochemistry* **35**, 2061-6 (1996).

25. Burgess, B. R., Dobson, R. C., Bailey, M. F., Atkinson, S. C., Griffin, M. D., Jameson, G. B., Parker, M. W., Gerrard, J. A. & Perugini, M. A. Structure and evolution of a novel dimeric enzyme from a clinically-important bacterial pathogen. *The Journal of biological chemistry* **283**, 27598-603 (2008).

26. Lusty, C. J. Catalytically active monomer and dimer forms of rat liver carbamoyl-phosphate synthetase. *Biochemistry* **20**, 3665-74 (1981).

27. Khayat, R., Batra, R., Bebernitz, G. A., Olson, M. W. & Tong, L. Characterization of the monomer-dimer equilibrium of human cytomegalovirus protease by kinetic methods. *Biochemistry* **43**, 316-22 (2004).

28. Batra, R., Khayat, R. & Tong, L. Molecular mechanism for dimerization to regulate the catalytic activity of human cytomegalovirus protease. *Nature structural biology* **8**, 810-7 (2001).

29. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods in enzymology* **276**, 307-26 (1997).

30. The CCP4 suite: programs for protein crystallography. *Acta crystallographica. Section D, Biological crystallography* **50**, 760-3 (1994).

31. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta crystallographica. Section D, Biological crystallography* **53**, 240-55 (1997).

32. Winn, M. D., Murshudov, G. N. & Papiz, M. Z. Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods in enzymology* **374**, 300-21 (2003).

33. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta crystallographica. Section D, Biological crystallography* **60**, 2126-32 (2004).

34. Devenish, S. R., Gerrard, J. A., Jameson, G. B. & Dobson, R. C. The high-resolution structure of dihydrodipicolinate synthase from *Escherichia coli* bound to its first substrate, pyruvate. *Acta crystallographica. Section F, Structural biology and crystallization communications* **64**, 1092-5 (2008).

35. Kefala, G., Evans, G. L., Griffin, M. D., Devenish, S. R., Pearce, F. G., Perugini, M. A., Gerrard, J. A., Weiss, M. S. & Dobson, R. C. Crystal structure and kinetic study of dihydrodipicolinate synthase from *Mycobacterium tuberculosis*. *The Journal of biological chemistry* **411**, 351-60 (2008).

36. Vaguine, A. A., Richelle, J. & Wodak, S. J. SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta crystallographica. Section D, Biological crystallography* **55**, 191-205 (1999).

37. Ramachandran, G. N. & Sasisekharan, V. Conformation of polypeptides and proteins. *Advances in protein chemistry* **23**, 283-438 (1968).

38. Phenix, C. P., Nienaber, K., Tam, P. H., Delbaere, L. T. & Palmer, D. R. Structural, functional and calorimetric investigation of MosA, a dihydrodipicolinate synthase from *Sinorhizobium meliloti* l5-30, does not support involvement in rhizopine biosynthesis. *Chembiochem: a European journal of chemical biology* **9**, 1591-602 (2008).

39. Blickling, S., Beisel, H. G., Bozic, D., Knablein, J., Laber, B. & Huber, R. Structure of dihydrodipicolinate synthase of *Nicotiana sylvestris* reveals novel quaternary structure. *Journal of molecular biology* **274**, 608-21 (1997).

40. Mirwaldt, C., Korndorfer, I. & Huber, R. The crystal structure of dihydrodipicolinate synthase from *Escherichia coli* at 2.5 Å resolution. *Journal of molecular biology* **246**, 227-39 (1995).

41. Girish, T. S., Sharma, E. & Gopal, B. Structural and functional characterization of *Staphylococcus aureus* dihydrodipicolinate synthase. *FEBS letters* **582**, 2923-30 (2008).

42. Blagova, E., Levdikov, V., Milioti, N., Fogg, M. J., Kalliomaa, A. K., Brannigan, J. A., Wilson, K. S. & Wilkinson, A. J. Crystal structure of dihydrodipicolinate synthase (BA3935) from *Bacillus anthracis* at 1.94 A resolution. *Proteins* **62**, 297-301 (2006).

43. Pearce, F. G., Perugini, M. A., McKerchar, H. J. & Gerrard, J. A. Dihydrodipicolinate synthase from *Thermotoga maritima*. *The Biochemical journal* **400**, 359-66 (2006).

44. Dobson, R. C., Griffin, M. D., Jameson, G. B. & Gerrard, J. A. The crystal structures of native and (*S*)-lysine-bound dihydrodipicolinate synthase from *Escherichia coli* with improved resolution show new features of biological significance. *Acta crystallographica. Section D, Biological crystallography* **61**, 1116-24 (2005).

45. Blickling, S., Renner, C., Laber, B., Pohlenz, H. D., Holak, T. A. & Huber, R. Reaction mechanism of *Escherichia coli* dihydrodipicolinate synthase investigated by X-ray crystallography and NMR spectroscopy. *Biochemistry* **36**, 24-33 (1997).

46. Pearce, F. G., Dobson, R. C., Weber, A., Lane, L. A., McCammon, M. G., Squire, M. A., Perugini, M. A., Jameson, G. B., Robinson, C. V. & Gerrard, J. A. Mutating the tight-dimer interface of dihydrodipicolinate synthase disrupts the enzyme quaternary structure: toward a monomeric enzyme. *Biochemistry* **47**, 12108-17 (2008).

Chapter Five

Summary and conclusions

5.1 Introduction

The focus of this work, as outlined in chapter one, was to use protein engineering to disrupt the quaternary structure of *M. tuberculosis* DHDPS and assess the impact of this disruption on enzyme function. This was undertaken in light of investigations with *E. coli* DHDPS, which showed the tetrameric quaternary structure to be essential for enzymatic activity and specificity,[1] and of the emerging paradigm in drug design of targeting protein-protein interfaces.[2] Chapter four presented the biophysical, kinetic and structural characterization of a dimeric variant of *M. tuberculosis* DHDPS, designed and produced by point mutation, as described in chapter three, and compared this protein with the tetrameric wild-type *M. tuberculosis* DHDPS, as detailed in chapter two. This comparison, taken together with recent research on DHDPS from other organisms, provides new insights into the relationship between oligomeric structure and catalytic function in DHDPS.

This chapter summarizes these results in the context of future drug design strategies and offers a rationale for the differing importance of the homotetrameric structure of *M. tuberculosis* DHDPS on enzyme function, in comparison with *E. coli* DHDPS. Through this understanding of the role of quaternary structure in DHDPS, new criteria are proposed to aid further investigation into drug design targeting DHDPS from various bacterial species.

5.2 The importance of quaternary structure

Chapter three (section 3.2) discussed general advantages that result from the association of subunits into higher order oligomers, such as functional gain, structural stability,[3] increased potential for allosteric regulation,[4] and decreased enzyme dynamics.[1]

5.2.1 The quaternary structure of wild-type *M. tuberculosis* DHDPS

The tetrameric quaternary structure of DHDPS is conserved in a broad range of species, as outlined in Table 1.4, chapter one, section 1.5.1. In chapter two, *M. tuberculosis* DHDPS was characterized and demonstrated unequivocally to be a stable tetramer in solution, consistent

with the crystal structure of the enzyme, deposited as PDB entry 1XXX. Wild-type *M. tuberculosis* DHDPS did not observably dissociate at high dilutions (~0.01 $mg.mL^{-1}$) in gel-filtration chromatography, or (0.06 $mg.mL^{-1}$) in analytical ultracentrifugation (AUC) experiments. Kinetic parameters comparable to *E. coli* DHDPS were determined using the coupled assay, as detailed in Table 2.5, chapter two, section 2.4.4, thus it seems wild-type *M. tuberculosis* DHDPS is functional as a non-dissociating tetramer.

X-ray crystallography data indicated that this tetramer was very similar to *E. coli* DHDPS, with near identical active-site geometry. *M. tuberculosis* DHDPS contains all active-site residues identified in *E. coli* DHDPS, except asparagine 248, as has been noted for other Gram-positive bacterial DHDPS orthologues.[5] The allosteric (*S*)-lysine binding site is absent from *M. tuberculosis* DHDPS, which is consistent with the lack of inhibition by (*S*)-lysine. Like the *E. coli* enzyme, *M. tuberculosis* DHDPS can be described as a dimer of "tight-dimers", although the "weak" interface between "tight-dimer" subunits had a higher buried surface area, with a larger number of interactions, than its *E. coli* counterpart (Figure 3.3, chapter three, section 3.4.1). This may contribute to the observation that a tyrosine mutation at the centre of the "weak" interface (L197Y) results in a dimeric version of *E. coli* DHDPS, but the corresponding mutation in *M. tuberculosis* DHDPS (A204Y) does not, as judged by BN-PAGE, Figure 3.7, chapter three, section 3.6.3.

5.2.2 The quaternary structure of DHDPS-A204R

The tetrameric quaternary structure was successfully disrupted in A204R *M. tuberculosis* DHDPS, as shown by BN-PAGE, gel-filtration chromatography and AUC. A dimeric quaternary structure was observed in the crystal structure of DHDPS-A204R; however, both monomeric and dimeric oligomeric states were observed with gel filtration and ultracentrifugation. The monomer-dimer equilibrium was quantified using sedimentation equilibrium AUC experiments, in 20 mM Tris.HCl, pH 8.0 150 mM NaCl yielding a dimerization dissociation constant ($K_D^{2\rightarrow1}$) of 51 μM (or 1.6 $mg.mL^{-1}$). This $K_D^{2\rightarrow1}$ is three orders of magnitude higher than the value of 33 nM determined for the natural dimer of *MRSA* DHDPS.[6] The CD spectra, reported in chapter three, were recorded at 0.2 $mg.mL^{-1}$ in the same buffer, and therefore the similarity between wild-type and DHDPS-A204R, predicted to be 85 % monomer by its $K_D^{2\rightarrow1}$ (using Equation 4.1, chapter four, section 4.4.2), indicates that the secondary structure has been preserved in the monomer of DHDPS-A204R.

X-ray crystallography shows the dimer of DHDPS-A204R to have analogous secondary and tertiary structure to wild-type *M. tuberculosis* DHDPS.

Pyruvate, the first substrate of DHDPS, was shown by both gel filtration and ultracentrifugation to shift the monomer-dimer equilibrium of DHDPS-A204R towards dimer. At pyruvate concentrations greater than or equal to 0.5 mM, the $K_D^{2\rightarrow1}$ shifted to below 0.01 mg.mL^{-1} (0.32 μM), as judged by analytical gel filtration (Figure 4.14, chapter four, section 4.4.3), and is only slightly affected by increasing pyruvate concentration. Similarly, pyruvate affects the $K_D^{2\rightarrow1}$ of *MRSA* DHDPS; although it caused only a 20-fold,[6] rather than a greater than 100-fold decrease. The lack of monomer observed for the interface mutants of *E. coli* DHDPS may be caused by the pyruvate analogue, α-ketoglutarate, found covalently bound in crystal structures,[1] as it is plausible that α-ketoglutarate may shift the monomer-dimer equilibrium in a similar manner to pyruvate.

The binding of pyruvate was shown to increase the thermal stability of DHDPS; irrespective of oligomeric state, as the apparent melting temperature was similarly increased (~7 °C) for both tetrameric wild-type enzyme and DHDPS-A204R (section 4.3.2, chapter four). The X-ray crystallographic structure of DHDPS-A204R showed pyruvate to be covalently bound to the key catalytic residue, lysine 171, forming polar contacts with backbone amides of two absolutely conserved threonines, as has been observed in other DHDPS enzymes (section 1.3.2, chapter one). One of these threonines (T54) is also a member of the proton relay, and is directly hydrogen bonded to Y117, which "reaches in" from the adjacent monomer. This provides a direct link between pyruvate and the tight-dimer interface.

The effect of pyruvate can be conceptualized in two different ways. Pyruvate could preferentially bind the dimer, stabilizing it thermodynamically and thus shift the equilibrium towards dimer **(Figure 5.1A)**. Alternatively, the monomer could be considered to exist in an ensemble of conformations,[7-9] and binding of pyruvate to the monomer could induce a conformation that preferentially forms dimers **(Figure 5.1B)**. Either or both models explain the experimental observations.

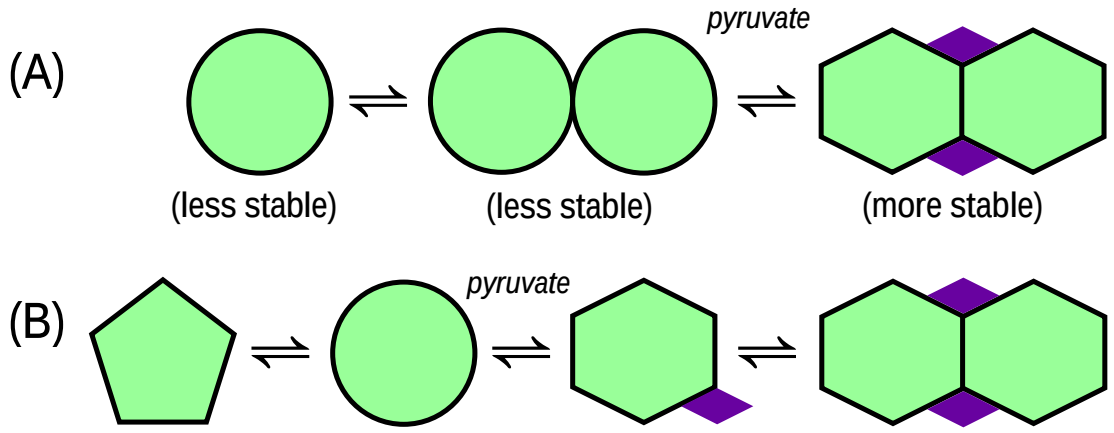


Figure 5.1: **Schematic representation of the two alternative explanations for the effect of pyruvate (purple diamond) on the monomer-dimer equilibrium of DHDPS-A204R. In (A) pyruvate binds to the dimer species increasing its stability. In (B) the monomer is conceptualized as adopting an ensemble of conformers in solution, pyruvate binds to the monomer species inducing a structural conformation that preferentially forms dimer.**

5.2.3 The quaternary structure of DHDPS-A204R is not critical for catalysis

Kinetic analysis showed DHDPS-A204R to have a turnover rate (k_{cat}) similar to wild-type *M. tuberculosis* DHDPS (as reported in chapter four, section 4.3.3). This was unexpected given that previous results with dimeric *E. coli* DHDPS showed a substantial drop in k_{cat} to 1.4 to 2.5 % that of wild-type *E. coli* DHDPS.[1] This decrease was in part due to the presence of a covalently bound inhibitor, α-ketoglutarate, believed to be acquired *in vivo* due to lower substrate specificity of the dimeric *E. coli* mutants.[1] Pyruvate was observed to displace α-ketoglutarate from the dimeric variants of *E. coli* DHDPS, as evident by mass spectrometry, although the activity was increased to only ~20 % that of wild-type.[1] DHDPS-A204R purified in the presence of pyruvate had a k_{cat} of ~85 % compared to wild-type *M. tuberculosis* DHDPS. Since wild-type dimeric *MRSA* DHDPS shows comparable activity to its tetrameric orthologues, such as *E. coli* DHDPS[6] and *M. tuberculosis* DHDPS, kinetic analysis would seem to indicate that DHDPS-A204R is more similar to the natural dimer of *MRSA* DHDPS than to the disrupted quaternary structure mutants of *E. coli* DHDPS.

Although the maximal catalytic rate (k_{cat}) was unaffected by disrupted quaternary structure, a decrease in substrate affinity is reflected in the increased Michaelis constant (K_M) of DHDPS-A204R for both pyruvate and (*S*)-ASA. The value of 0.33 (±0.03) mM for pyruvate is

increased two-fold in comparison to wild-type *M. tuberculosis* DHDPS, and three-fold compared to the wild-type *MRSA* DHDPS. A greater increase of ~10-fold was seen in the disrupted quaternary structure mutants of the *E. coli* enzyme, such as dimeric DHDPS-L197Y, and the monomer-dimer-tetramer DHDPS-Y107W.[1,10] However, in these cases the K_M values are likely to reflect the presence of the covalent adduct, α-ketoglutarate, in the active site,[1,10] whereas this adduct is absent in *M. tuberculosis* DHDPS-A204R, as shown by the crystal structure (Figure 4.16B, chapter four, section 4.5.3), and therefore the K_M values are not directly comparable.

Monomeric DHDPS-A204R, missing the active-site residue, Y117, is predicted to have reduced activity, consistent with the concentration-dependent decrease in activity of *MRSA* DHDPS reported in literature and attributed to its dissociation into monomer.[6] The percentage of monomer under the assay conditions for *M. tuberculosis* DHDPS-A204R has not been determined, since the monomer-dimer equilibrium was quantified in the absence of pyruvate with sedimentation equilibrium AUC experiments. In an analogous system, the different kinetic parameters for alternative oligomeric forms of porphobilinogen synthase manifest as double hyperbolae in the kinetic analysis,[11] and these were suggested as being a general feature of enzymes existing in multiple oligomeric states.[9] The single hyperbola observed for DHDPS-A204R, which exhibited a good fit to the models used, seems to suggest that a single oligomeric species was characterized, or that the kinetics of the DHDPS-A204R monomer and dimer are equivalent.[12] However, it should be noted that to successfully fit the kinetic data of DHDPS-A204R to double hyperbolae knowledge of the percentage of monomer under assay conditions would be required, which was not available in this case.

In summary, the k_{cat} of DHDPS-A204R clearly indicates that tetrameric quaternary structure is not critical for catalysis in *M. tuberculosis* DHDPS. The k_{cat} was only slightly affected by disrupted quaternary structure, and was similar to other wild-type DHDPS enzymes, including the naturally dimeric *MRSA* DHDPS. The increased K_M values for both substrates, indicate that the dimeric species is somewhat less efficient than the wild-type enzyme, with k_{cat}/K_M values of 360 (±52) and 108 (±15) $s^{-1}.mM^{-1}$ for pyruvate and (*S*)-ASA, respectively. However, this effect is not substantial enough to justify targeting inhibitors to disrupt the weak interface of *M. tuberculosis* DHDPS, which precludes interface disruption as a means of inhibiting this enzyme.

5.2.4 The quaternary structure of DHDPS-A204R may decrease stability

The tetrameric quaternary structure of *M. tuberculosis* DHDPS may be important for the stability of the enzyme. The disrupted quaternary structure of DHDPS-A204R resulted in decreased thermal stability shown by the substantial drop in the apparent melting temperature (T_m^{app}) of ~30 ºC compared to wild-type, as shown by both DSF and the coupled assay (section 4.2.1 and 4.3.2, chapter four). However, the change in T_m^{app} needs to be considered cautiously, since DHDPS-A204R has been engineered to expose a hydrophobic interface that would ordinarily be buried, which undoubtedly contributes to the loss of thermal stability. The mutant DHDPS-A204R also seemed prone to aggregation in the absence of pyruvate, probably due to dissociation into monomers, which hinted that quaternary structure may play a role in protecting against aggregation.

5.2.5 Comparison of the tight-dimer interfaces of characterized DHDPS enzymes from different species

The modulated importance of tetramerization in catalysis for *M. tuberculosis* DHDPS, compared with *E. coli* DHDPS, suggested similarities with the naturally dimeric *MRSA* DHDPS. The tight-dimer of *MRSA* DHDPS had an increased number of contacts and greater buried surface area than *E. coli* DHDPS, and this was proposed to reduce the dynamics in a manner similar to tetramerization in *E. coli* DHDPS.[6] Since the X-ray crystal structures of various DHDPS had been determined, their tight-dimer interfaces were examined in comparison with *MRSA* DHDPS,[6] and this analysis is extended here to provide a rationale for the range of activities observed for dimeric DHDPS enzymes.

The association of monomers to form the dimer of *M. tuberculosis* DHDPS-A204R buries 1414 (±3) Å^2 of the solvent-accessible surface area.[13] This is within the 1200-2000 Å^2 range not expected to result in large conformational changes upon complex formation,[14] and thus is consistent with the unchanged secondary structure of DHDPS-A204R monomer compared to the wild-type tetramer (discussed in section 5.2.2). The interface involves 40 (±1) residues, seven of which are completely buried, including, Y117, and its hydrophobic stacking partner, Y116. The substrate binding residues T54 and R148 form inter-subunit hydrogen bonds with Y117, which together contribute four interface hydrogen bonding interactions to the total of 14 observed in all three dimers in the asymmetric unit. There were only slight differences with the equivalent tight-dimer interfaces in the crystal structure of wild-type *M. tuberculosis*

DHDPS, burying a surface area of 1466 (±5) Å^2, involving the same 40 interface residues as DHDPS-A204R, and with 16 inter-subunit hydrogen bonds apparent in all four tight-dimers in the asymmetric unit.

An aromatic stacking interaction between two tyrosines, Y90 and its counterpart on the opposite subunit was identified at the tight-dimer interface of *M. tuberculosis* DHDPS **(Figure 5.2A)**, but is absent in the *E. coli* enzyme **(Figure 5.2B)**. The energetically favourable interactions of buried or partially buried aromatic residues contribute to the stability of tertiary and quaternary structure,[15] and aromatic clusters have been linked with the improved thermal stability in enzymes from thermophiles.[16] Y90 contributes to the hydrophobic stack involving Y116 and Y117, resulting in six tyrosines forming hydrophobic and π interactions, with the T-shaped and parallel-displaced geometry common in proteins.[15,17] The larger network of interacting aromatic side chains of *M. tuberculosis* DHDPS-A204R in comparison to *E. coli* DHDPS-L197D **(Figure 5.2)** is proposed to decrease the dynamic fluctuations, such as subunits reorientating themselves in relation to each other, and consequently maintain a catalytic rate comparable with the wild-type enzyme.

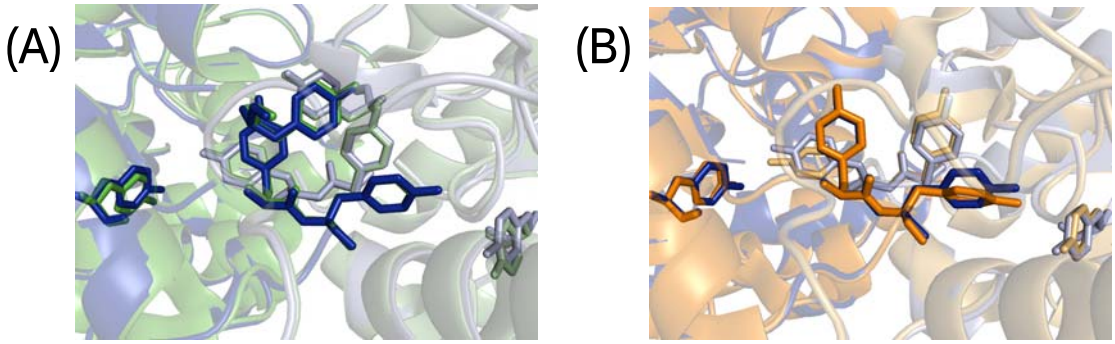


Figure 5.2: **The aromatic network buried in the tight-dimer interface of (A) wild-type *M. tuberculosis* DHDPS (green) overlaid with mutant DHDPS-A204R (blue) and (B) wild-type *E. coli* DHDPS (orange) overlaid with mutant DHDPS-L197Y (blue). Chain B is indicated by lighter colours. The interface tyrosines and proton-relay tyrosines from both subunits are shown as thick lines.**

One of the aromatic residues, Y117 in *M. tuberculosis* DHDPS, is of particular interest, being the key catalytic residue that interdigitates across the tight-dimer interface **(Figure 5.2A)**. This provides a rationale for dynamic motion affecting enzyme activity (chapter one, section 1.5.3). In *M. tuberculosis* DHDPS, this residue (*M. tuberculosis*: Y117, *E. coli*: Y107,

MRSA: Y109) is twisted relative to the equivalent residue in the *E. coli* enzyme (chapter two, section 2.6.3). In *MRSA* DHDPS, the twisting of this tyrosine in comparison to the *E. coli* enzyme is postulated to provide closer aromatic stacking interactions than in *E. coli* DHDPS.[6]

The aromatic cluster at the tight-dimer interface in *MRSA* DHDPS involved four tyrosines, analogous to *E. coli* DHDPS, so it is unlikely the aromatic network decreased dynamic motions in a manner similar to that postulated for *M. tuberculosis* DHDPS-A204R. However, analysis of the tight-dimer interface in *MRSA* DHDPS identified a salt bridge involving an arginine (R115) and glutamate (E275). Salt bridges have also been linked with the improved thermal stability in enzymes from thermophiles.[18] The salt bridges in naturally dimeric *MRSA* DHDPS are thus hypothesized to decrease movement, in a similar manner to the aromatic cluster in *M. tuberculosis* DHDPS-A204R.

The largest and smallest tight-dimer interfaces determined for DHDPS to date have been from *MRSA* and *E. coli* DHDPS, respectively.[6] These interfaces, along with the previously unexamined interface of *E. coli* DHDPS-L197Y were analyzed for comparison with the dimer interface of *M. tuberculosis* DHDPS-A204R **(Table 5.1)**. A similar number of hydrogen bonds was identified in the tight-dimer interfaces of *MRSA* and *M. tuberculosis* DHDPS-A204R, which was more than double the number identified for *E. coli* DHDPS-L197Y **(Table 5.1)**.

Table 5.1: Comparison of the tight-dimer interface in dimeric DHDPS and their tetrameric counterparts, prepared using the web-based PISA program.[13]

Species	Enzyme	PDB code	SISA[a] due to interface		H-bonds formed	Interface residues
			(Å^2)	% of total		
E. coli	DHDPS-L197Y	2OJP	1315 (±1)	11.2 %	7	36 (±0.0)
E. coli	wild-type DHDPS	1YXC	1289 (±7)	11.2 %	7	38 (±0.7)
M. tuberculosis	DHDPS-A204R	*pending*	1414 (±3)	12.0 %	14	40 (±0.8)
M. tuberculosis	wild-type DHDPS	1XXX	1466 (±5)	12.5 %	16	40 (±0.4)
MRSA	wild-type DHDPS	3DAQ	1661 (±17)	13.2 %	17	49 (±1.2)

[a] SISA = solvent-inaccessible surface area

When considering the dimeric DHDPS variants, *E. coli* DHDPS-L197Y, *M. tuberculosis* DHDPS-A204R and *MRSA* DHDPS, there is a trend of increasing interface size, participating

residues and number of hydrogen bonds **(Table 5.1** and **Figure 5.3)**. This aligns well with their catalytic ability, as *E. coli* DHDPS-L197Y shows a low rate of catalysis and low substrate affinity, *M. tuberculosis* DHDPS-A204R shows a high catalytic rate and somewhat reduced substrate affinity, and *MRSA* DHDPS has a high rate of catalysis and high substrate affinity.

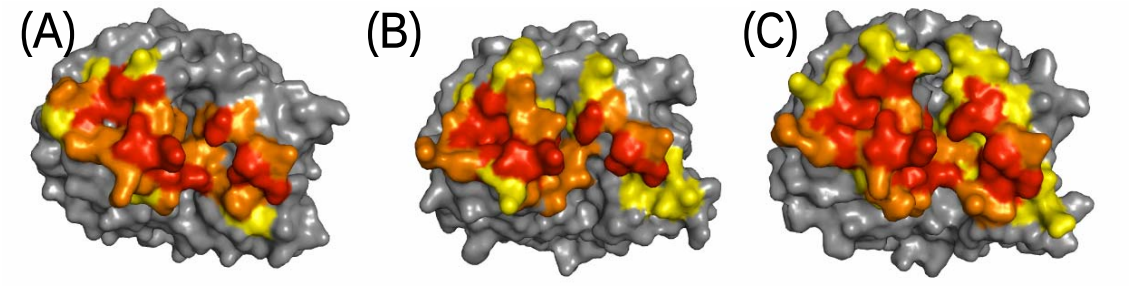


Figure 5.3: **Surface view of the monomer showing the tight-dimer interface from (A) the dimeric *E. coli* mutant DHDPS-L197Y, (B) the dimeric *M. tuberculosis* mutant DHDPS-A204R, and (C) wild-type dimeric *MRSA* DHDPS. The residues contributing to the interface, shown in various shades depending on their buried surface area (5 - 44 % yellow, 45 - 84 % orange, ≥85 % red) as defined by PISA.[13]**

Previous investigations with homodimers found a lack of correlation between overall binding energy and interface size,[19] probably because interface size is just one component that contributes to the binding energy of dimerization. However, a correlation between interface size and strength in DHDPS enzymes seems to be reflected in the dissociation constant $K_D^{2\rightarrow1}$ determined in the absence of pyruvate, as the $K_D^{2\rightarrow1}$ of *M. tuberculosis* DHDPS-A204R is orders of magnitude higher than the value determined for *MRSA* DHDPS (mentioned in section 5.2.2). This analysis is entirely consistent with the hypothesis that increased dynamics are associated with decreased catalytic ability, and correlated with the strength of the interface, as proposed by Burgess *et al.*[6]

5.3 Suggestions for future work

This work focused on disruption of the weak interface to create a dimeric variant of *M. tuberculosis* DHDPS; however, the observation that DHDPS-A204R exists in monomer-dimer equilibrium, suggests a monomeric mutant of *M. tuberculosis* DHDPS could

also be created to further probe the role of quaternary structure. Additionally, the determination of the $K_D^{2\rightarrow1}$ of DHDPS-A204R in the presence of pyruvate, with sedimentation equilibrium AUC experiments, may allow the kinetic characterization of the monomeric species of DHDPS-A204R. The lower concentration needed to investigate this $K_D^{2\rightarrow1}$ may require fluorescent-labelled protein and the recent instrumental advancement of coupling fluorescence detection with AUC, used in the study of *MRSA* DHDPS.[6]

The interesting effects of the substrate, pyruvate, on the thermal stability and oligomerization of DHDPS-A204R has relevance to developing the understanding of the relationship between dynamics and oligomerization in DHDPS. Consequently, further studies are suggested monitoring the effect of pyruvate and temperature on secondary structure using CD spectrometry. The crystal structure without pyruvate bound, or of the monomeric species of DHDPS-A204R may also be informative when considering the two alternative hypotheses for the effect of pyruvate on oligomerization as outlined in **Figure 5.1**.

5.4 Conclusion

Remarkably, the weak interface between the tight-dimer subunits of *M. tuberculosis* DHDPS was disrupted by a single site-directed mutation to create a functioning dimeric species, DHDPS-A204R. The catalytic competency of this enzyme can be understood by comparison to other species, specifically by comparing the extent of the tight-dimer interactions; that is, only DHDPS enzymes with a poorly developed tight-dimer interface require the tetrameric structure for catalysis. In the case of *M. tuberculosis* DHDPS, the homotetrameric quaternary structure perhaps plays a role in maintaining protein stability, or maybe the association of tight-dimers into the tetramer involves vestigial interfaces which are no longer required for function.

New criteria should be considered in the future when exploring interface disruption of DHDPS as an approach to drug design. The tight-dimer interface should be analyzed and compared with those of *E. coli*, *M. tuberculosis* and *MRSA* DHDPS, and those DHDPS orthologues with poorly developed tight-dimer interfaces should be pursued, as it is more likely these will exhibit a similar loss of catalytic rate due to disruption of quaternary structure as that observed with *E. coli* DHDPS.

For *M. tuberculosis* DHDPS, future drug discovery efforts will be directed at the active site, rather than targeting protein-protein interfaces. However, this work provides a framework through which to discriminate DHDPS enzymes that are likely to be inactive as dimers, such that future drug discovery efforts using the interface targeting approach can be directed there.

5.5 References

1. Griffin, M. D., Dobson, R. C., Pearce, F. G., Antonio, L., Whitten, A. E., Liew, C. K., Mackay, J. P., Trewhella, J., Jameson, G. B., Perugini, M. A. & Gerrard, J. A. Evolution of quaternary structure in a homotetrameric enzyme. *Journal of molecular biology* **380**, 691-703 (2008).

2. Gerrard, J. A., Hutton, C. A. & Perugini, M. A. Inhibiting protein-protein interactions as an emerging paradigm for drug discovery. *Mini reviews in medicinal chemistry* **7**, 151-7 (2007).

3. Mei, G., Di Venere, A., Rosato, N. & Finazzi-Agro, A. The importance of being dimeric. *The FEBS journal* **272**, 16-27 (2005).

4. Traut, T. W. Dissociation of enzyme oligomers: a mechanism for allosteric regulation. *Critical reviews in biochemistry and molecular biology* **29**, 125-63 (1994).

5. Blickling, S., Renner, C., Laber, B., Pohlenz, H. D., Holak, T. A. & Huber, R. Reaction mechanism of *Escherichia coli* dihydrodipicolinate synthase investigated by X-ray crystallography and NMR spectroscopy. *Biochemistry* **36**, 24-33 (1997).

6. Burgess, B. R., Dobson, R. C., Bailey, M. F., Atkinson, S. C., Griffin, M. D., Jameson, G. B., Parker, M. W., Gerrard, J. A. & Perugini, M. A. Structure and evolution of a novel dimeric enzyme from a clinically-important bacterial pathogen. *The Journal of biological chemistry* **283**, 27598-603 (2008).

7. Ma, B., Shatsky, M., Wolfson, H. J. & Nussinov, R. Multiple diverse ligands binding at a single protein site: a matter of pre-existing populations. *Protein science* **11**, 184-97 (2002).

8. James, L. C. & Tawfik, D. S. Conformational diversity and protein evolution--a 60-year-old hypothesis revisited. *Trends in biochemical science* **28**, 361-8 (2003).

9. Jaffe, E. K. Morpheeins--a new structural paradigm for allosteric regulation. *Trends in biochemical science* **30**, 490-7 (2005).

10. Pearce, F. G., Dobson, R. C., Weber, A., Lane, L. A., McCammon, M. G., Squire, M. A., Perugini, M. A., Jameson, G. B., Robinson, C. V. & Gerrard, J. A. Mutating the tight-dimer interface of dihydrodipicolinate synthase disrupts the enzyme quaternary structure: toward a monomeric enzyme. *Biochemistry* **47**, 12108-17 (2008).

11. Tang, L., Stith, L. & Jaffe, E. K. Substrate-induced interconversion of protein quaternary structure isoforms. *The Journal of biological chemistry* **280**, 15786-93 (2005).

12. Kacser, H., Sauro, H. M. & Acerenza, L. Enzyme-enzyme interactions and control analysis. 1. The case of non-additivity: monomer-oligomer associations. *European journal of biochemistry* **187**, 481-91 (1990).

13. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *Journal of molecular biology* **372**, 774-97 (2007).

14. Lo Conte, L., Chothia, C. & Janin, J. The atomic structure of protein-protein recognition sites. *Journal of molecular biology* **285**, 2177-98 (1999).

15. Burley, S. K. & Petsko, G. A. Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science* **229**, 23-8 (1985).

16. Kannan, N. & Vishveshwara, S. Aromatic clusters: a determinant of thermal stability of thermophilic proteins. *Protein engineering* **13**, 753-61 (2000).

17. McGaughey, G. B., Gagne, M. & Rappe, A. K. π-Stacking interactions. Alive and well in proteins. *The Journal of biological chemistry* **273**, 15458-63 (1998).

18. Bjork, A., Mantzilas, D., Sirevag, R. & Eijsink, V. G. Electrostatic interactions across the dimer-dimer interface contribute to the pH-dependent stability of a tetrameric malate dehydrogenase. *FEBS Letters* **553**, 423-6 (2003).

19. Brooijmans, N., Sharp, K. A. & Kuntz, I. D. Stability of macromolecular complexes. *Proteins* **48**, 645-53 (2002).

Chapter Six

Experimental

6.1 Materials and equipment

Unless otherwise stated, chemicals were purchased from Sigma-Aldrich Chemical Company Ltd. (Castle Hill, Australia) and all enzyme manipulations occurred at 4 °C or on ice. SDS-PAGE, native PAGE gels, associated buffers, gel boxes, and power units were purchased from Invitrogen (Auckland, New Zealand). Protein ladders were purchased from Fermentas (*via* Global Sciences, Christchurch, N.Z.), Sigma-Aldrich, and Invitrogen. Bio-Rad protein assay kit and DNA ladders were supplied by Bio-Rad Laboratories (Auckland, N.Z.) Restriction enzymes were supplied by Roche (Christchurch, N.Z.). *E. coli* strains and mutagenesis kits were purchased from Stratagene (*via* Global Sciences, Christchurch, N.Z.). Column chromatography media were purchased from GE Healthcare (Auckland, N.Z.) as pre-packed columns or loose beads to be packed in XK type columns.

Centrifugation at 4 ºC was mainly carried out in an Eppendorf Centrifuge model 5810R, using different Eppendorf rotors for small volumes (≤1.5 mL, F-45-30-11, max speed of 14000 rpm), moderate volumes (≤50 mL, F-45-30-11, max speed of 12000 rpm), and large volumes (≤500 mL, A-4-81, max speed of 4000 rpm). Chromatography columns were run using a UPC-900 ÄKTA FPLC (GE Healthcare), which was loaded using either a Gilson Minipuls M312 peristaltic pump or a syringe. The fractions were collected with a Frac-950 fraction collector (GE Healthcare). Ultraviolet (UV) absorbance was recorded either by using a Diode Array spectrophotometer (Hewlett Packard, model 8452A or 8453), a Nanodrop ND 1000 spectrophotometer (Thermo Scientific) or a Smart Spec Plus spectrophotometer (Bio-Rad).

6.2 Microbiological and molecular biological methods

The experimental section from the PhD thesis "Why is dihydrodipicolinate synthase a tetramer?", authored by Griffin,[1] was an invaluable reference for the development of experimental methods. The laboratory text "Molecular Cloning: A laboratory manual", authored by Sambrook *et al.*,[2] provided an excellent resource for protocols for DNA and bacterial strain manipulation, as did instruction manuals from Stratagene and Novagen.

All bacterial cultures were grown under sterile conditions. All media and equipment were sterilized by either autoclaving at 121 ºC for 20 min or filtration through a 0.2 μm syringe tip filter from Millipore (*via* Biolab, Christchurch, N.Z.), or were purchased sterile. All manipulations of bacterial cultures occurred within the sterile environment created by a flame or a laminar fume hood. Standard sterile technique was employed and when necessary appropriate controls were performed alongside experiments to monitor for possible contamination.

6.2.1 Bacterial strains

In this study, several bacterial strains were explored for optimum protein expression. These strains were originally derived from *E. coli* BL21 (DE3), and all derivatives except BL21 (DE3) pGroESL were purchased directly from Stratagene. BL21 (DE3) pGroESL was produced by transforming competent cells of BL21 (DE3) from Stratagene with pGroESL. The plasmid pGroESL, which expresses the *E. coli* folding chaperone proteins GroES and GroEL, was kindly donated by Celia Webby (Massey University, N.Z.). The plasmid pGroESL was originally developed by George Lorimer's Lab (E. I. Dupont De Nemours and Company, Wilmington, DE, U.S.A.) and has been shown to increase soluble protein yield.[3] The bacterial strains used in this research are shown in **Table 6.1**.

Table 6.1: *E. coli* strains used plasmid manipulation and protein expression.

Strain	Derived from	Key features	Antibiotic^r
XL1-Blue	K-12	Cloning, plasmid preps, supports growth of vectors carrying mutations, allows for blue/white screening on X-gal plates	Tet^r
DH5α	K-12	Cloning, plasmid preps	--
BL21 (DE3) pLysS	BL21 (DE3)	Lacks *Lon* & *ompT* proteases, inducible & supressed basal expression (with T7 lysozyme) of T7 RNA polymerase	Cam^r
BL21 (DE3) pGroESL	BL21 (DE3)	Lacks *Lon* & *ompT* proteases, inducible expression of T7 RNA polymerase & expression of heat shock folding chaperonins GroEL & GroES	Cam^r
BL21-CodonPlus® (DE3) RIL	BL21 (DE3)	Lacks *Lon* & *ompT* proteases, inducible expression of T7 RNA polymerase & expression of tRNAs recognizing rare codons (AGA/AGG, AUA, and CUA)	Cam^r
BL21-CodonPlus® (DE3) RP	BL21 (DE3)	Lacks *Lon* & *ompT* proteases, inducible expression of T7 RNA polymerase & expression of tRNAs recognizing rare codons (AGG/AGA and CCC)	Cam^r

Additional *E. coli* strains used were XL1-Blue and DH5α, also obtained from Stratagene. Both strains are recombination (*rec*A) deficient, improving plasmid insert stability, do not produce endonuclease I (*end*A), which degrades DNA, and thus are excellent for mutagenesis and plasmid preparation. The genotypes of these and other strains used are listed in **Table 6.2**.

Table 6.2: *E. coli* strains used and their associated genotypes.

Strain	Genotype
XL1-Blue	*recA*1 *endA*1 *gyrA*96 thi-1 *hsdR*17 *supE*44 *relA*1 *lac*$^-$ [F' *proAB*$^+$ *lacI*^{q} *lacZ*ΔM15 (Tet^r)]
DH5α	*supE*44 ΔlacU169 (Φ80 *lacZ*ΔM15) *hsdR*17 *recA*1 *endA*1 *gyrA*96 *thi*-1 *relA*1
BL21 (DE3)	F$^-$ o*mpT hsdS*(r_B- m_B-) *dcm*$^+$ *gal λ*(DE3)
BL21 (DE3) pLysS	F$^-$ *ompT hsdS*(r_B- m_B-) *dcm*$^+$ *gal λ*(DE3) [pLysS Cam^r]
BL21 (DE3) pGroESL	F$^-$ *ompT hsdS*(r_B- m_B-) *dcm*$^+$ *gal λ*(DE3) [pGroESL *groES groEL* Cam^r]
BL21-CodonPlus® (DE3) RIL	F$^-$ *ompT hsdS*(r_B- m_B-) *dcm*$^+$ Tet^r *gal λ*(DE3) *endA* Hte [*argU ileY leuW* Cam^r]
BL21-CodonPlus® (DE3) RP	F$^-$ *ompT hsdS*(r_B- m_B-) *dcm*$^+$ Tet^r *gal λ*(DE3) *endA* Hte [*argU proL* Cam^r]

6.2.2 Plasmids

The plasmids utilized in this work were based on several different vector systems, as listed in **Table 6.3**. The genes *dap*A and *dap*B from *E. coli* had been previously cloned into the Stratagene plasmid pBluescript KS+ by others.[4-8] pTM1520, containing *dap*B from *T. maritima* with a purification tag (MGSDKIHHHHHH) at its amino terminus, was a generous gift from Scott Lesley and Heath Klock (Joint Center for Structural Genomics, Genomics Institute of the Novartis Research Foundation, San Diego, USA). *E. coli* XL1-Blue strains were available that had been transformed with these plasmids.

The plasmid containing *dap*A from *M. tuberculosis* was a generous gift from collaborators Manfred S. Weiss and Georgia Kefala (EMBL Hamburg Outstation, Hamburg, Germany). The plasmid pMTB02 consisted of *M. tuberculosis dap*A cloned into pETM-11. The vector pETM-11 is a modified version of the Novagen plasmid pET-24d containing an N-terminal polyhistidine tag with a TEV protease cleavage site **(Appendix A)**. The pET system allows for inducible protein expression in which T7 RNA polymerase transcribes genes cloned under control of a T7*lac* promoter **(**detailed in **Appendix A)**. Thus protein expression in this system requires both a bacterial strain with produces T7 RNA polymerase, such as *E. coli* BL21 (DE3),[9] and an inducer, such as isopropyl thiogalactoside (IPTG), to bind to the *lac* repressor and allow gene transcription.[10]

The plasmid pMTB02 was introduced into *E. coli* BL21 (DE3) pLysS and BL21-CodonPlus® (DE3) RP using competent cells prepared by the calcium chloride method,[1] and into BL21 (DE3) pGroESL using the electroporation method.[1] Successful transformants were identified by the conferred kanamycin resistance (Kan^r).

Table 6.3: Plasmids used for protein expression and their associated genotypes.

Plasmid	Ref.	Derived from	Reference/source	Relevant genotype
pJG001	5	pBluescript KS+	(Stratagene)	:: *dap*A (*E. coli*), Amp^r
pJK001	4	pBluescript KS+	(Stratagene)	:: *dap*B (*E. coli*), Amp^r
pTM1520	11	pMH1	12	:: *dap*B (*T. maritima*), Amp^r
pTOPO-TM1520	(in-house)	pTOPO	11	:: *dap*B (*T. maritima*), Amp^r
pMTB02	13	pETM-11	14	:: *dap*A (*M. tuberculosis*), Kan^r
pMTBA204Y	(this work)	pETM-11	14	:: *dap*A-ala204tyr (*M. tuberculosis*), Kan^r
pMTBA204D	(this work)	pETM-11	14	:: *dap*A-ala204asp (*M. tuberculosis*), Kan^r
pMTBA204R	(this work)	pETM-11	14	:: *dap*A-ala204-arg (*M. tuberculosis*), Kan^r
pGroESL	3	pACYC184	15	:: *groES* and *groEL* (*E. coli*), Cam^r
pRK793	16	pMal-C2	(New England Biolabs)	:: NIa protease-ser-219-val (*TEV*), Amp^r

The insertion of the *dap*B gene from pTM1520 into the pTOPO vector by Dr Grant Pearce resulted in a new plasmid called pTOPO-TM1520.[11] *E. coli* BL21 (DE3) pLysS strains that had been transformed by pTOPO-TM1520 were available.

The plasmid pRK793 was acquired through the non-profit organization Addgene (http://www.addgene.org), which distributes certain useful plasmids found in the published literature. Kapust and colleagues rationally designed a tobacco etch virus (TEV) protease mutant, S219V, which is 100-fold more stable than wild-type protease and does not self-cleave.[16] The plasmid pRK793 expresses the mutant TEV protease with a polyhistidine tag (GHHHHHHH) and a polyarginine tag, on the N- and C-terminus respectively, for purification and solubility purposes.[16] *E. coli* BL21-CodonPlus (DE3)® RIL strains that had been transformed by pRK793 were available.

6.2.3 Antibiotics and other media supplements

Stock and working concentrations of antibiotics used for bacterial selection are shown in **Table 6.4**. Stock solutions were prepared with the appropriate solvent, sterilized by filtration

and stored at -20 ºC. For selection of bacterial strains, the appropriate antibiotics were added to cultures, using a 1/1000 dilution, to give the final working concentration.

Table 6.4: Antibiotic concentrations used for bacterial selection.

Antibiotic	Abbreviation	Solvent	[Stock] mg.mL^{-1}	[Working] µg.mL^{-1}
Ampicillin	Amp	d-H_2O	100	100
Chloramphenicol	Cam	EtOH	30	30
Kanamycin	Kan	d-H_2O	30	30
Tetracycline	Tet	MeOH	15	15

Stock solutions of IPTG and arabinose were prepared for inducing protein expression. To suppress protein expression in inducible systems prior to induction, glucose was added to overnight starter cultures.[17] IPTG was also used, along with X-gal, for blue/white screening, which is a method for determining mutagenesis efficiency.[18] Stock solutions were prepared with the appropriate solvent, as listed in **Table 6.5**, sterilized by filtration and stored at -20 ºC.

Table 6.5: Nutritional supplements and other additives concentrations.

Additive/supplement	Abbreviation	Solvent	[Stock] mg.mL^{-1}	[Working] µg.mL^{-1}
Arabinose	--	d-H_2O	10	100
meso-Diaminopimelate	DAP	d-H_2O	10	50
Glucose	--	d-H_2O	50	500
Isopropyl thiogalactoside	IPTG	d-H_2O	60	60
5-Bromo-4-chloro-3-indolyl-β-D-galactopyranoside	X-gal	DMF	20	--

6.2.4 Media and plate preparation

Luria-Bertani medium (LB)

LB broth components were supplied in a ready to use mix, in powder form. The LB was dissolved in distilled water with a ratio of 20 g to 1 L of water and autoclaved to sterilize. Flasks of 0.75 and 1 L were prepared and autoclaved, as well as vials of 3, 5 and 10 mL.

Super optimal broth (SOB), with catabolite repression (SOC)

A stock solution of 250 mM potassium chloride (KCl) was prepared in advance and sterilized in an autoclave. For catabolite repression, a 1 M glucose stock solution was also prepared in advance, sterilized by filtration and stored at -20 ºC.

To prepare SOB, 10 g of tryptone, 2.5 g of yeast extract and 0.25 g NaCl were dissolved in 480 mL of distilled water. Before making up to 500 mL with distilled water, 10 mL of 250 mM KCl was added and the pH adjusted to 7.0 with concentrated NaOH. The SOB medium was aliquoted and autoclaved.

For SOC media, 1 M glucose was added to SOB, under sterile conditions just prior to use, to give a final concentration of 20 mM glucose.

ZYM-5052 auto-inducing medium

Based on the protocol developed by Studier,[17] several stock solutions were prepared in advance and stored at room temperature. For the 50× 5052 stock, 125 mL of glycerol, 12.5 g of glucose and 50 g of lactose were dissolved in 500 mL of distilled water, aliquoted into 15 mL and autoclaved. For the 50× M stock, 167.5 g of disodium phosphate heptahydrate ($Na_2HPO_4{\cdot}7H_2O$), 85.1 g of monopotassium phosphate (KH_2PO_4), 66.9 g of ammonium chloride (NH_4Cl), and 17.8 g of sodium sulfate (Na_2SO_4) were dissolved in 500 mL of distilled water, aliquoted into 15 mL and autoclaved. For the 1000× trace metal solution, a stock solution of 0.1 M iron trichloride ($FeCl_3$) was dissolved in 1 % concentrated HCl, and subsequently diluted and combined with stock solutions of other metals to give 50 mM $FeCl_3$, 20 mM calcium chloride ($CaCl_2$), 10 mM manganese dichloride ($MnCl_2$), 10 mM zinc sulfate ($ZnSO_4$), 2 mM cobalt dichloride ($CoCl_2$), 2 mM copper dichloride ($CuCl_2$), 2 mM nickel chloride ($NiCl_2$), 2 mM sodium molybdate (Na_2MoO_4), 2 mM sodium selenite (Na_2SeO_3), 2 mM boric acid (H_3BO_3) in approximately 60 mM HCl. For the 1 M magnesium sulfate ($MgSO_4$) stock, 6.02 g of $MgSO_4$ was dissolved in 50 mL of distilled water and filter sterilized.

In a flask, 7.5 g of tryptone and 3.75 g of yeast extract were dissolved in 750 mL of distilled water and autoclaved. Just prior to inoculation and supplementation with appropriate

antibiotics, 15 mL of 5052 stock, 15 mL of M stock, 150 μL of trace metal stock, and 750 μL of 1 M $MgSO_4$ were added to the flask, under sterile conditions.

Agar plates

Bacteriological agar was dissolved in LB medium at a ratio of 15 g to 1 L and autoclaved to sterilize. Once cooled or reheated to ~50 ºC, the appropriate antibiotics were added and the molten media were poured onto sterile Petri dishes in the Bio-hazard hood. Any bubbles were removed by flaming the surface before the agar set. For blue/white screening, plates were pre-treated by spreading 14 μL of 60 $mg.mL^{-1}$ IPTG and 40 μL of 20 $mg.mL^{-1}$ X-gal on to the surface, and allowing 30 min for the DMF to evaporate.

6.2.5 Bacterial cultures

Agar plates containing appropriate selective antibiotics were streaked with a bacterial strain from a glycerol freeze stock, an overnight culture or a fresh single colony, using a flame-sterilized nichrome wire loop. The plates were incubated for at least 9 hrs at 37 ºC. Individual colonies were selected using a sterile pipette tip and used to inoculate 3 or 5 mL of liquid medium containing appropriate selective antibiotics. Starter cultures were grown overnight at 37 ºC, with shaking at 200 rpm, and subsequently used to inoculate larger quantities of media.

Glucose was included with starter cultures when protein production was under the control of the T7*lac* promoter, preventing low level protein expression.[17] Auto-inducing media worked on the same principle, with induction repressed by glucose until it was sufficiently depleted, and with the addition of many components to support the growth of high density culture.[17] Bacterial cultures grown in LB media required the addition of the inducer IPTG (or arabinose, in the case of pTM1520), after sufficient bacterial growth (OD_{600} of ~0.60), in order to induce protein expression.

6.2.6 Bacterial strain storage

Glycerol freezer stocks were used to store bacterial strains. These were prepared by centrifuging (5000 rpm, 3 min, 4 ºC), in a screw-top cryo-storage tube, an aliquot (1.5 mL) from an overnight culture (3.0 mL) produced by inoculation with a single bacterial colony. The supernatant was removed and the cell pellet was gently re-suspended in 0.5 mL of overnight culture and 0.5 mL of sterile 30 % glycerol, and then stored at -80 ºC.

6.2.7 Plasmid preparation by alkaline lysis

Based on the protocol outlined in Sambrook *et al.*,[19] the following stocks were prepared:

Solution #1, pH 8.0	50 mM glucose
(stored at -20 ºC)	25 mM Tris.HCl
	10 mM EDTA
Solution #3	3 M potassium acetate
(stored at -20 ºC)	11.5 % (v/v) glacial acetic acid
TE Buffer, pH 8.0	40 mM Tris.HCl
(stored at room temp.)	1 mM EDTA

Small scale preparation

A single bacterial colony was selected from an agar plate and used to inoculate a vial of LB broth, containing appropriate antibiotics, and incubated overnight (37 ºC, 200 rpm). Subsequently, the overnight culture was centrifuged (7000 rpm, 2 min, 4 ºC) in two Eppendorf tubes (1.5 mL/tube). Solutions #1 and #3 were defrosted and kept cold on ice. Solution #2 was freshly prepared:

Solution #2	0.2 M NaOH
(kept at room temp.)	1 % (w/v) SDS

The supernatant was removed from the Eppendorf tubes by aspiration, and the pellet was re-suspended in 100 μL of ice-cold solution #1 with vigorous vortexing and chilled on ice. After 5 min, 200 μL of freshly prepared solution #2 was added to each tube, mixed by gentle inversion, and the tubes were chilled on ice. After 10 min, 150 μL of ice-cold solution #3 was added to the preparations, mixed by flicking and inverting until a white precipitate formed, and once again chilled on ice (for 10 min). The preparations were centrifuged (12000 rpm, 5 min, 4 ºC) and the supernatants (400 μL/tube) were carefully transferred to another set of Eppendorf tubes and the pellets were discarded.

The DNA was precipitated by the addition of an equal volume (400 μL/tube) of isopropanol, mixed by vortexing, and allowed to stand at room temperature for 2 min. After centrifugation (12000 rpm, 5 min, 4 ºC), the pellets were rinsed with 70 % ethanol (1 mL) and air-dried for at least 10 min, before being dissolved in TE buffer (50 μL).

For further purification, RNA was digested by adding 5 μL of DNAse-free RNAse (Roche) and incubating at 37 ºC for at least 30 min. RNAse and other impurities were removed either by a phenol/chloroform extraction step or using spin-columns from a QIAquick gel extraction kit (Biolab Ltd, Auckland, New Zealand). Plasmid identity, purity and concentration were examined by restriction digests mapped with agarose gel electrophoresis (as outlined in sections 6.2.9 and 6.2.10, respectively). The concentration and purity were determined more quantitatively by using a 1/100 dilution in a quartz cuvette and reading the absorbance at 260 and 280 nm, as outlined in Sambrook *et al*.[2]

Phenol/chloroform protein extraction

After incubation with DNAse-free RNAse, another 350 μL of TE buffer and 400 μL of ice cold solution of 25:24:1 phenol:chloroform:isoamyl alcohol was added, mixed with vortexing, and centrifuged (12000 rpm, 5 min, 4 ºC). The upper (aqueous) layer (200 μL) was then transferred to a new Eppendorf tube, discarding the lower (organic) layer, and 400 μL of ice cold solution of 24:1 chloroform:isoamyl alcohol was added. This process repeated several times. DNA was precipitated from the upper layer by the addition of isopropanol (500 μL), as before, rinsed with 70 % ethanol, air dried, and dissolved in 50 μL TE buffer.

Spin-column DNA cleanup

Spin-columns contain a silica gel membrane that absorbs DNA in the presence of high salt and neutral pH, while impurities are washed through. The method was carried out as described in the QIAquick® Spin Handbook,[20] using buffers QG, PB and EB supplied in the accompanying kit. To the DNA sample, 3 volumes of QG buffer (165 μL) and an equal volume of isopropanol (55 μL) were added. The mixture was loaded into the upper compartment of the spin-column and centrifuged for 1 min (at 14500 rpm) to bind DNA. After washing with PB buffer (0.75 mL), the DNA was eluted from the column with EB buffer (10 mM Tris.HCl, pH 8.5) and subjected to centrifugation (14500 rpm, 1 min, 20 ºC).

Large scale/high purity preparation

When sequencing quality DNA preparations were required, the procedure described above was scaled up by doubling the volumes of all stock solutions. A 5 mL vial of overnight culture was harvested and re-suspended in 200 μL of ice-cold solution #1, in a 10 mL Falcon tube. DNA was extracted as previously described, except additions of 400 μL and 300 μL

were made of solution #2 and #3, respectively. After centrifugation, the supernatant was retained and DNA was precipitated by adding isopropanol (800 µL). The mixture was centrifuged again, and the pellet retained and rinsed with 70 % ethanol, air dried, and dissolved in 50 µL TE buffer. Subsequent plasmid purification steps were as outlined for small scale preparations.

6.2.9 Restriction digests of plasmids

Restriction digests were performed in 10 µL volumes. Typically ~1 µg of plasmid DNA (2.5 to 5 µL) was digested with 1 µL of the appropriate restriction enzyme and specified restriction buffer (10× concentration stock). For pGroESL a double digest with *Eco*R I and *Hin*d III was performed. For pMTB02 and variants **(**map of restriction sites in **Appendix A)**, a single digest with *Xho* I/*Nco* I and a double digest with *Xho* I and *Nco* I were performed. These were incubated at 37 ºC for 2 hours before being resolved with electrophoresis.

A typical experimental setup for pMTB02 double digest is as follows:

Double digest	5 µL plasmid DNA
	1 µL buffer H
	1 µL *Xho* I
	1 µL *Nco* I
	2 µL d-H_2O

6.2.10 Agarose gel electrophoresis

This protocol was based on method outlined in Sambrook *et al.*:[21]

6× Loading buffer	30 % (v/v) glycerol
(stored at -20 ºC)	0.25 % (w/v) bromophenol blue
	0.25 % (w/v) xylene cyanol
1× TAE buffer, pH 8.0	4 mM Tris.acetate
(stored at 4 ºC)	1 mM EDTA

Restriction fragments were resolved on a 1 % agarose gel against Hyperladder I (Bioline) or Read-Load™ DNA ladder (Bio-Rad), which provide standards ranging from 200 to 10000 bp, or 100 to 12000 bp, respectively. A solution of 1 % agarose (0.3 g agarose dissolved in 30 mL

of 1× TAE buffer) was heated and allowed to cool slightly (~50 ºC) before being poured into a gel casting tray, followed by well comb insertion. The gel was allowed to set completely (~30 min) before the comb was removed and transferred into a gel tank containing cold 1× TAE buffer (stored at 4 ºC). The DNA ladder (5 µL for Hyperladder I, 2.5 µL for Read-Load™ DNA ladder) was loaded with 5 µL of loading buffer. Digested or uncut DNA (5 µL) was loaded with an equivalent amount of loading buffer (5 µL). Electrophoresis was conducted at 80/90 V for 90 min or until the bromophenol band was nearing the end of the gel. The gel was stained with ethidium bromide (0.5 $\mu g.mL^{-1}$) for 20 min. The stained DNA fragments were visualized under 302 nm UV radiation and photographed using the Chemi Genius2 Bio Imaging System (Syngene).

6.2.11 PCR site-directed mutagenesis

The method was carried out as described in the Stratagene QuikChange site-directed mutagenesis manual.[18] The advantage of the QuikChange site-directed mutagenesis system is that it allows double stranded plasmid DNA, such as that obtained by alkaline lysis, to be used as template DNA.[18]

Primer design

Mutagenic primers introduce site-specific mutations by being incorporated into the newly synthesized DNA during the polymerase chain reaction (PCR) step. For each mutation, two primers are required to anneal to both the forward and reverse of the same sequence found on opposing strands of template DNA. The primers were designed following guidelines in the Stratagene QuikChange site-directed mutagenesis manual.[18]

Primers were between 25 and 45 bases in length, with the desired mutation in the middle flanked by ~10 to 15 bases of correct sequence on both sides. The melting temperatures (T_m) were ≥78 ºC as determined by **Equation 6.1**.

$$T_m = 81.5 + 0.41 \times (\%\mathrm{GC}) - 675 / N - \%mismatch \qquad \textbf{Equation 6.1}\ [18]$$

Here T_m is calculated from N, the length of the primer, the percentage of bases mismatched to the template DNA, and the GC content. Optimally, primers should be composed of 40 % or more G and C bases and terminate with either G or C.

A typical example of primer design is as follows:

To introduce a mutation from alanine (A) to arginine (R) at position 204, the *dap*A sequence from pMTB02 was examined **(Figure 6.1)**.

```
                                        204
Parent strand          5′-aac ctg ccc tgg ctg gcc atg ggc gcc acg ggc-3′
Primer #1                5′-C CTG CCC TGG CTG CGC ATG GGC GCC ACG-3′
Primer# 2                3′-G GAC GGG ACC GAC GCG TAC CCG CGG TGC-5′
```

Figure 6.1: The DHDPS-A204R primers with the sequence alteration in bold blue and with the codon for amino-acid position 204 underlined.

Possible codons for arginine are CGA, CGC, CGG and CGT, only one of which shares a base with the codon GCC used for A204, thus CGC was chosen to minimize the mismatch between the primer and the parent strand. Taking thirteen bases of correct sequence to flank either side resulted in forward (#1) and complementary reverse (#2) primers with melting temperatures of 78.2 ºC, as calculated by **Equation 6.1**. Both primers terminated in two C or G bases and had a GC content of 70 %.

Other mutagenic primers were designed in an analogous way (listed in chapter three, Table 3.2) and purchased from Invitrogen.

PCR reaction conditions

For each experiment, two PCR reactions were prepared per set of mutagenic primers (100 ng.μL^{-1}), using 5 and 50 ng of template DNA (as described in **Table 6.6**), along with a control reaction. The sample reaction used pMTB02 isolated by a high purity preparation (as described in section 6.2.7) from *E. coli* DH5α pMTB02, further purified by spin-column technology.

The control reaction was performed using pWhitescript with the mutagenic primers provided in the Stratagene QuikChange site-directed mutagenesis kit. The pWhitescript plasmid contains a point mutation that generates a stop codon in the middle of *lacZ* gene, disrupting the synthesis of β-galactosidase.[18] If the mutagenesis reaction is successful, the stop codon is removed, restoring enzyme activity and giving bacterial colonies a blue phenotype on plates pre-treated with IPTG and X-gal for blue/white screening.[18] Thus, the mutational efficiency

can be determined by the proportion of blue versus white colonies resulting from the control reaction, prepared as outlined in **Table 6.6**:

Table 6.6: Preparation for site-directed mutagenesis.

	Control (µL)	5 ng of Sample (µL)	50 ng of Sample (µL)
10× Reaction buffer	5.0	5.0	5.0
Forward primer (#1)	1.25	1.25	1.25
Reverse primer (#2)	1.25	1.25	1.25
dNTP mix	1.0	1.0	1.0
Plasmid	2.0	1.0	10.0
d-H_2O	39.5	40.5	31.5
Total	50.0	50.0	50.0

The final addition to the 50 µL reaction mixture was 1 µL of *PfuTurbo*® DNA polymerase. This enzyme replicates both plasmid strains with high fidelity without displacing the mutagenic primers.[18] The reactions were performed in thin-walled PCR tubes, using a thermocycler equipped with a "hot top". The cycling parameters are summarised in **Table 6.7**. After temperature cycling, the amplification reactions were incubated on ice for 2 min.

Table 6.7: PCR cycling parameters for site-directed mutagenesis.

Segment	Cycles	Temperature (ºC)	Time (min)
1	1	95	0.5
2	16	95	0.5
		50	1.0
		68	13.0

Template digestion

Endonuclease enzyme, *Dpn* I (1 µL), was added to each of the amplification reactions. This enzyme is specific for methylated and hemimethylated DNA, and hence digests the non-mutated parent template DNA but leaves the mutation-containing amplified DNA intact. The enzymatic digestion reaction was gently mixed with a pipette and incubated at 37 ºC for at least 1 hour. Agarose gel electrophoresis (as described in section 6.2.10) was used to confirm the presence of amplified DNA in the *Dpn* I treated reactions, prior to it being used to transform *E. coli* XL1-Blue supercompetent cells, as described in the following section.

6.2.12 Transformation of *E. coli* strains with plasmids

Transforming XL1-Blue supercompetent cells

Frozen *E. coli* XL1-Blue supercompetent cells (purchased from Stratagene) were defrosted on ice and aliquoted in 50 µL volumes into sterile pre-chilled Eppendorf tubes. To the separate aliquots, 2 µL of 0.142 M β-mercaptoethanol and 2 µL of the *Dpn* I treated DNA was added and mixed by swirling, followed by incubation on ice for 30 min. A transformation control was also prepared, containing 1 µL of pUC18 (0.1 $ng.\mu L^{-1}$). The transformation reactions were heat shocked for 45 s at 42 ºC, then immediately put on ice. After 2 min, 900 µL of SOC was added to each reaction and cells were allowed to express antibiotic resistance by incubation at 37 ºC for at least 1 hour.

Selection of transformants and monitoring transformation efficiency

For each plasmid, two different types of plate were prepared, one containing antibiotics selective for the strain, and another containing antibiotic selective for the strain and plasmid.

For the pWhitescript control for site-directed mutagenesis, the plates were pre-treated with IPTG and X-gal for blue/white screening (as outlined in section 6.2.4). On the appropriate plates, 200 µL of SOC, followed by 10 µL of transformed cells, were pipetted, then spread out with a glass spreader, sterilized by ethanol and flame.

For the plasmid of interest, 100 µL of transformed competent cells was also pipetted and spread on plates containing appropriate antibiotics. The remaining transformed cell mixture was centrifuged at 14000 rpm for 1 min, the cell pellet re-suspended in 100 µL of SOC and spread on the same type of plates.

All plates were incubated overnight at 37 ºC. Site-directed mutagenesis efficiency with pWhitescript was calculated from the number of blue bacterial colonies divided by the total number of colonies on the agar plate and was reported as a percentage. The introduced mutation was confirmed by sequencing (section 6.2.13) of plasmid preparations (section 6.2.7) from the newly transformed XL1-Blue cells.

Competent cell preparation by calcium chloride

A vial of LB medium, containing appropriate antibiotics, was inoculated with bacteria from a freshly streaked agar plate and grown overnight in a 37 ºC shaker. A small flask of LB medium (50 mL) was inoculated with 1 mL of starter culture and allowed to grow in a 37 ºC shaker until an OD_{600} of 0.8 was reached (~1.25 hrs). After cooling on ice for 10 minutes, cells were harvested by centrifugation (4000 rpm, 10 min, 4 ºC), under sterile conditions, the supernatant was removed and the pellet re-suspended, with gentle pipetting, in cold sterilized 10 mM $CaCl_2$ (10 mL). Again, the cells were centrifuged (4000 rpm, 10 min, 4 ºC), and the supernatant was removed. The same volume (10 mL) of cold sterilized $CaCl_2$ at 10× the concentration (100 mM), was used to re-suspend the cells. The supernatant was removed from the cells for the third time by centrifugation, and the cells were re-suspended in cold 100 mM $CaCl_2$ (2 mL). The competent cells were pipetted into sterile Eppendorf tubes in 200 µL aliquots and stored on ice for less than five hours before transformation.

Transformation for calcium chloride method

To each aliquot of calcium chloride competent cells, 0.1 ng of plasmid DNA (dissolved in either d-H_2O/TE buffer) was added, followed by careful mixing, by swirling or stirring with the pipette tip. Two transformation controls were also prepared, one containing no plasmid DNA and another containing 1 µL of pUC18 (0.1 ng.μL^{-1}). The competent cells were treated in an analogous manner to that previously described for XL1-Blue supercompetent cells.

Competent cell preparation for the electroporation method

LB medium was inoculated using an overnight culture produced by inoculation with a single bacterial colony. The culture was allowed to grow in a 37 ºC shaker until it reached an OD_{600} of 0.5 and then put on ice for 30 min. The cells were harvested by centrifugation at 4 ºC for 5 min at 4000 rpm, the supernatant was removed and the cells re-suspended in one culture volume of cold sterilized d-H_2O. This process was repeated twice, using 0.5 and 0.05 culture volumes, before final re-suspension in 0.05 culture volume of 10 % glycerol. The competent cells were pipetted into sterile Eppendorf tubes in 40 µL aliquots and stored at -80 ºC.

Transformation by electroporation

To each aliquot of competent cells was added 0.1 ng of plasmid DNA (dissolved in either d-H_2O or TE buffer), with careful mixing by swirling or stirring with the pipette tip. The cells

were transferred to sterilized (with ethanol and UV), cooled and dried electroporation glass cuvettes and electroporated using a Gene Pulser (BioRad), set to 2.5 kV and 25 μF at 200 Ω. Immediately afterwards, the cells were suspended in 1 mL of SOC, and then pipetted into sterile Eppendorf tubes. The competent cells were incubated at 37 ºC for 1 hour then transformed cells were selected as previously outlined.

6.2.13 DNA Sequencing

DNA sequencing was performed at Canterbury Sequencing, part of the School of Biological Sciences at the University of Canterbury. The plasmids were sequenced using a capillary ABI3100 Genetic Analyzer, (Applied Biosystems Inc.) using BigDye® Terminator and T7 primers with a procedure based on the dideoxynucleic acid chain termination method, developed by Sanger.[22]

6.3 Biochemistry general methods

Unless otherwise stated, all enzymes were manipulated on ice or 4 ºC and buffered at pH 8.0 by 20 mM Tris.HCl. pH measurements were carried out using an Ultra Basic UB-10 pH/mV meter (Denver Instruments). Proteins were concentrated using spin columns from Vivascience (via Global Sciences, Christchurch, N.Z.) The laboratory text "Guide to protein purification", edited by Deutscher,[23] provided an excellent resource for biochemistry techniques and protein purification, as did various editions of "Protein purification", authored by Scopes,[24,25] and "Enzyme assays: a practical approach", authored by Eisenthal and Danson.[26]

6.3.1 Determining protein concentration

Bradford Assay

The colorimetric assay, originally developed by Bradford,[27] utilizes the shift in absorbance maximum, from 465 nm to 595 nm, of Coomassie brilliant blue G-250 dye upon protein binding, in acidic conditions, to determine protein concentration.[28] An 800 μL volume of protein, appropriately diluted by distilled water, was placed in a disposable polystyrene cuvette. After the addition of 200 μL of Protein Assay Dye Reagent (Bio-Rad), containing dye, phosphoric acid and methanol, the assay was incubated at room temperature for 10 minutes, before measuring absorbance, using water as a blank.

As an improvement to the original method developed by Zor & Selinger,[29] the ratio of absorbances at 590 nm and 450 nm, was recorded, giving a linear relationship for protein concentration from 0.2 to 20 $\mu g.mL^{-1}$. Protein concentrations were determined based on a calibration curve generated using known concentrations of bovine serum albumin (BSA). All measurements were carried out in triplicate. Different proteins bind Coomassie blue dye slightly differently, depending on their amino-acid sequence; thus the values determined were relative rather than actual concentration values.[28]

Ultraviolet (UV) absorbance methods

Proteins absorb light at 280 nm due to aromatic amino acids, such as tryptophan and tyrosine, hence there is high-level of variability for absorbance at this wavelength between proteins.[28] This method only accurately determines concentration for homogenous solutions of a protein with known amino-acid sequence. The relationship between protein concentration (c) and absorbance (A) across path-length (l) is described by Beer's law as outlined in **Equation 6.2**:

$$A = \varepsilon \cdot c \cdot l \qquad \textbf{Equation 6.2}$$

The extinction coefficient (ε) for absorbance at 280 nm by specific proteins was predicted based on amino-acid sequence using the web-based program "Protein Calculator".[30] This allowed for direct determination of protein concentration by reading absorbance at 280 nm across a 1 mm path-length using a Nanodrop spectrophotometer blanked with buffer.

6.3.2 Sodium dodecyl sulfate polyacrylamide gel electrophoresis

Protein samples were routinely analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), using the NuPAGE® Bis-Tris system developed by Invitrogen, which is similar to the Tris-Glycine system,[31] except it operates at neutral pH resulting in better protein stability and resolution.[32] Concentrated NuPAGE® (4×) sample and (20×) running buffer were purchased or made in-house, and gave final concentrations as follows:

1× MOPS running buffer, pH 7.7 (stored at 4 ºC)	50 mM 3-morpholinopropane-1-sulfonic acid (MOPS) 50 mM Tris.base 0.1 % (w/v) SDS 1 mM EDTA

1× LDS sample buffer, pH 8.5	10 % (v/v) glycerol
(stored at -20 ºC)	0.106 M Tris.HCl
	0.141 M Tris.base
	2% (w/v) lithium dodecyl sulfate (LDS)
	0.51 mM EDTA
	0.22 mM SERVA® Blue G-450
	0.175 mM phenol Red

The NuPAGE® MOPS running buffer (× 20) was prepared in-house by dissolving 209.2 g of MOPS, 121.2 g of Tris.base, 20 g of SDS, and 3 g of EDTA in 750 mL of distilled water to make 1 L of buffer.[32]

Proteins were resolved on pre-cast NuPAGE® 4-12 % Bis-Tris gradient polyacrylamide gels (Invitrogen) against protein standards, such as PageRuler™ (Fermentas), Novex® Sharp Pre-stained (Invitrogen) or Wide Range SigmaMarker™ (Sigma-Aldrich). Samples were prepared as outlined in **Table 6.8**, using various dilutions to give 10-20 µg of protein per sample for optimal resolution.

Table 6.8: Sample preparation for SDS-PAGE.

	Volume added (µL) for 1/10 dilution	Volume added (µL) for 1/5 dilution	Volume added (µL) for 1/2 dilution
d-H_2O	5.5	4.5	1.5
4× Loading buffer	2.5	2.5	2.5
10× Reducing agent	1.0	1.0	1.0
Protein	1.0	2.0	5.0
Total	10.0	10.0	10.0

The samples were heated in boiling water for 10 min, centrifuged briefly, then loaded on the Bis-Tris gel secured in an XCell *SureLock*™ Mini-Cell gel box (Invitrogen) adequately filled with NuPAGE® MOPS running buffer (1×) and with 500 µL of NuPAGE® antioxidant pipetted into the upper chamber.

Electrophoresis was conducted at 180 V at room temperature for 60 min or until the dye band neared the bottom of the gel. The gel was subsequently removed from the plastic casing and stained with Coomassie blue dye for 30 min to an hour, with shaking, at ~55 rpm. Standard one-dimensional PAGE staining protocols were used and required the preparation of the following solutions:[31]

Coomassie blue stain	0.1 % (w/v) Coomassie brilliant blue G-250
(stored at room temp.)	50 % (v/v) methanol
	10 % (v/v) glacial acetic acid
De-stain	5 % (v/v) methanol
(stored at room temp.)	10 % (v/v) glacial acetic acid

The gels were de-stained overnight, or for several hours in microwave heated de-stain. The stained protein bands were visible to the naked eye and recorded using the Chemi Genius2 Bio Imaging System (Syngene).

6.3.3 Blue-native polyacrylamide gel electrophoresis

BN-PAGE was carried out using a protocol modified from that originally developed by Schagger *et al.*:[33]

Anode buffer, pH 8	100 mM Tris.HCl
Cathode buffer, pH 8	50 mM Tricine
(stored at 4 ºC)	30 mM Tris.HCl
	0.02 % (w/v) Coomassie brilliant blue G-250
4× Sample treatment buffer, pH 8	200 mM Tris.HCl
(stored at -20 ºC)	1.4 % (w/v) Coomassie brilliant blue G-250
	20 % (v/v) glycerol

Blue-native polyacrylamide gel electrophoresis (BN-PAGE) was performed with pre-cast NativePAGE 4-16 % Bis-Tris gradient polyacrylamide gels (Invitrogen). Protein samples were prepared to a final volume of 20 μL and contained 5 μL of sample treatment buffer (4×). The samples were loaded directly on to the gel, secured in an XCell *SureLock*™ Mini-Cell gel box (Invitrogen), with blue cathode running buffer in the upper chamber and colourless anode running buffer in the lower chamber. Electrophoresis was conducted at 100 V at room

temperature for 2 hours or until the Coomassie blue dye had run off the end of the gel, evident from colouration in the anode buffer. Gels were de-stained as described in section 6.3.2. On occasion, gels were also fixed prior to de-staining, but this was found to be unnecessary.

Fixing solution	50 % (v/v) methanol
(stored at room temp.)	10 % (v/v) glacial acetic acid

6.3.4 Preparation of dialysis tubing

Dialysis tubing (cellulose membrane, MW cut-off >12000 kDa) was prepared according to the manufacturer's instructions by rinsing with continuously flowing distilled water for 1 hour or more, followed by 2 min submerged in a solution of 0.3 % (w/v) sodium sulfide, heated to 80 ºC, before being soaked in freshly boiled distilled water. After 3 min, this water was removed and replaced with ~5 L of water and 10 mL of concentrated sulfuric acid. After 3 min in the 0.2 % (v/v) sulfuric acid, the solution was poured off and the tubing washed twice more by submerging in recently boiled distilled water. After the water cooled, the tubing was removed and stored in 0.1 % (w/v) sodium azide at room temperature.

6.4 Optimized over-expression and purification of wild-type *M. tuberculosis* DHDPS and variants

Buffer A	20 mM Tris.HCl, pH 8.0
	250 mM NaCl
	10 mM pyruvate
	5 % (v/v) glycerol
	2 mM β-mercaptoethanol (β-ME)
	10 mM imidazole
Buffer B	20 mM Tris.HCl, pH 8.0
	1 M NaCl
	10 mM pyruvate
	5 % (v/v) glycerol
	2 mM β-ME
	10 mM imidazole

Buffer C	20 mM Tris.HCl, pH 8.0 250 mM NaCl 10 mM pyruvate 5 % (v/v) glycerol 2 mM β-ME 50 mM imidazole
Buffer D	20 mM Tris.HCl, pH 8.0 250 mM NaCl 10 mM pyruvate 5 % (v/v) glycerol 2 mM β-ME 250 mM imidazole
Storage buffer	20 mM Tris.HCl, pH 8.0 250 mM NaCl 5 % (v/v) glycerol 2 mM β-ME *with/without* 10 mM pyruvate

The protocol has been optimized from original method developed by Kefala *et al*.[13]

6.4.1 Growth of *E. coli* BL21 (DE3) pGroESL

A glycerol freezer stock of *E. coli* BL21 (DE3) pGroESL containing the relevant DHDPS expression plasmid was streaked out on an LB agar plate containing kanamycin and chloramphenicol and incubated for ≥9 hours at 37 ºC. Single colonies were used to inoculate 5 mL of LB broth, containing kanamycin, chloramphenicol and glucose. The starter cultures were grown overnight at 37 ºC with shaking (200 rpm). Six flasks containing 750 mL of either LB broth or auto-inducing (ZYM-5052) media (see section 6.2.4) were supplemented with 750 μL of kanamycin and chloramphenicol and inoculated with 1.5 mL of starter culture. These large cultures were grown at 37 ºC with shaking for ~4 hours or until an OD_{600} of 0.60.

For LB cultures, over-expression was induced with 750 μL of IPTG (60 $mg.mL^{-1}$, 0.25 M). After induction the cells were grown overnight at room temperature. For auto-inducing

cultures, the flasks were moved to a room temperature shaker after the initial growth period and grown for an additional 12 hours. The cells were then chilled on ice for 30 minutes, harvested by centrifugation (4000 rpm, 15 min, 4 ºC), and washed by re-suspension in cold buffer A, followed by further centrifugation (8000 rpm, 10 min, 4 ºC). The supernatant was discarded and the cell pellets were re-suspended in a small amount of buffer A. Typically, one 4 L preparation yielded 14-16 g of cells using auto-inducing media and 8-12 g of cells using LB broth. Several cell preparations were accumulated and stored at -80 ºC before further processing.

6.4.2 Preparation of crude extract

Several cell pellets (~70 g) were thawed and re-suspended in ~200 mL of buffer A. Crude cell-free lysates were produced by ultrasonication for 5 min at 4 °C in 20 s pulses with 4 min delays (~80 % amplitude), using a Sonics Vibracell sonicator. Following sonication, the cell extract was clarified by centrifugation (10000 rpm, 30 min, 4 ºC). The supernatant was collected and further clarified by centrifugation (10000 rpm, 60 min, 4 ºC). A 500 µL aliquot of the final supernatant was reserved for analysis.

6.4.3 Affinity chromatography

Purification using His_6-tag affinity chromatography involved several steps, including washing with three different buffers, which was found to improve protein purity. The supernatant from centrifugation was loaded onto a His_6-tag affinity column (2× 5 mL HisTrap™ Crude or HisTrap™ FF, GE Healthcare) pre-equilibrated with buffer A, and subsequently washed with 5 column volumes of the same buffer. This was followed by 5 column volumes each of buffer B, buffer C and buffer D. Washing with buffer B increased the concentration of NaCl from 0.25 to 1 M, eluting proteins associating with the column resin by ionic interactions. Buffer C returned the NaCl concentration to 0.25 M and increased imidazole from 10 to 50 mM, eluting proteins weakly associating with the Ni^{2+} column resin. Buffer D increased imidazole to 250 mM to elute the His_6-tagged *M. tuberculosis* DHDPS, which was strongly associating with the resin. The peak fractions, as determined by absorbance at 280 nm, were tested for activity using the *o*-aminobenzaldehyde assay (section 6.4.4). Active fractions were pooled and 500 µL was set aside for further analysis, prior to buffer exchange into buffer A (section 6.4.5).

6.4.4 *o*-Aminobenzaldehyde assay

The colorimetric *o*-aminobenzaldehyde assay is a semi-quantitative method for monitoring DHDPS activity,[34] and has been developed for high throughput in our lab. It was carried out in a 96 well plate with 5 µL of test solution presumed to contain DHDPS and 160 or 100 µL of assay mixture prepared as described in **Table 6.9**. The 50× preparation can be stored at 4 ºC for several days (Dr Sean Devenish, *pers. comm.*).

Table 6.9: The *o*-aminobenzaldehyde assay mixture for DHDPS activity.

	Volume added (µL) for 1× preparation	Volume added (µL) for 50× preparation
200 mM Tris.HCl, 40 mM pyruvate, pH 8.0	150	7500
400 mM *o*-aminobenzaldehyde (in EtOH)	5	250
~100 mM (*S*)-ASA	5	250
Total added to well	160	100

The assay was incubated at 37 ºC for 20 min. The colour change due to DHDPS activity was developed by the addition of 100 µL of 10 % (w/v) trichloroacetic acid (TCA), giving deep purple colouration to the active fractions. Negative and positive controls were also prepared, with no DHDPS and DHDPS from a previous purification, respectively.

6.4.5 Buffer exchange

Two methods were used to change the buffer of protein solutions.

Dialysis

The pooled fractions were transferred into rinsed dialysis tubing (see section 6.3.4) and clamped either end. The protein solution was dialyzed overnight against 100 volumes of cold buffer, at 4 ºC, with stirring.

Desalting column

The pooled fractions were loaded in appropriate volumes onto a desalting column (HiPrep^{TM} 26/10 or HiTrap^{TM} Desalting, GE Healthcare), pre-equilibrated with buffer. The protein was eluted with the required volume of buffer, and its corresponding peak was collected based on absorbance at 280 nm or the *o*-aminobenzaldehyde assay.

6.4.6 His_6-tag cleavage

In order to cleave His_6-tag, protein fractions were incubated overnight in TEV protease (purchased or purified as outlined in 6.6) in the presence of 5 mM ethylenediaminetetraacetic acid (EDTA) and 1 mM dithiothreitol (DTT).[16,35] After SDS-PAGE confirmed the cleavage of the tag**,** the cleaved DHDPS was passed through a His_6-tag affinity column, pre-equilibrated with buffer A, to remove the His_6-tag peptide fragment. Active fractions as determined by the *o*-aminobenzaldehyde assay (described in section 6.4.4) were pooled and 500 µL was set aside for further analysis.

6.4.7 Gel-filtration chromatography

The pooled fractions were concentrated using an ultrafiltration spin column (Vivaspin 15, MW cut-off >10000 kDa, volume 2-8 mL, Vivascience) in preparation for gel filtration. The concentrate was loaded onto a gel-filtration column (HiLoad Superdex, 16/60, GE Healthcare) pre-equilibrated with storage buffer to remove any remaining impurities and the active peak (as determined by the *o*-aminobenzaldehyde assay) was collected.

Mutant *M. tuberculosis* DHDPS-A204R was stored in buffer containing 10 mM pyruvate, whereas wild-type was stored in buffer without pyruvate and kept at 4 or -20 °C for several weeks without detectable degradation. Aliquots of 500 µL taken at different purification steps were analyzed by SDS-PAGE (section 6.3.2), the Bradford assay (section 6.3.1), and the coupled assay (as described in sections 6.9.2 and 6.9.4).

6.4.8 Optimization trials

Several parameters were varied during optimization trials to increase soluble protein yield.

Extraction buffers

In early trials, full scale protein preparations (4.5 L) were grown in LB media supplemented with kanamycin and chloramphenicol, at 37 °C with shaking, until an OD_{600} of ~0.50 was reached. Cells were harvested and ultrasonicated for 25 min at 4 ºC in 2 s pulses with 10 s delays (~30 % amplitude), in each of the different extraction buffers listed in **Table 6.10**. The cell extract was clarified by centrifugation and His_6-tagged protein was purified in an analogous manner to that described in sections 6.4.3 and 6.4.5.

Table 6.10: Utilizing different extraction buffers for purification.

Strain	Extraction buffer		
	Type	Imidazole	Other components
BL21 (DE3) pLysS	Tris.HCl, pH 8.0	0 mM	NaCl, glycerol, β-ME
BL21 (DE3) pLysS	Tris.HCl, pH 8.0	10 mM	NaCl, glycerol, β-ME
BL21 (DE3) pLysS	Phosphate, pH 7.4	5 mM	NaCl, glycerol
BL21 (DE3) pLysS	Tris.HCl, pH 8.0	10 mM	NaCl, glycerol, β-ME, pyruvate

Bacterial strains

For subsequent trials, smaller scale preparations (1 L) of LB media containing kanamycin and chloramphenicol were grown at 20 ºC for ≥12 hours when they reached an OD_{600} of ~0.50. The smaller culture size allowed for simultaneous growth of three different bacterial strains, BL21 (DE3) pGroESL, BL21 (DE3) pLysS, and BL21-CodonPlus® (DE3) RP, all of which contained pMTB02. After induction with IPTG, the cultures in each of the three different flasks were grown at 20 ºC for 20 hours, and then harvested by centrifugation. After re-suspension in buffer A, cells were lysed by ultrasonication for 3 min at 4 ºC in 2 s pulses with 10 s delays (~30 % amplitude), and 500 µL was set aside for further analysis. Insoluble proteins were removed by centrifugation (11000 rpm, 40 min, 4 ºC). The total and soluble protein fractions for each bacterial strain were examined by SDS-PAGE and the coupled assay.

Lysis procedures

The effect of different lysis procedures was also examined with smaller scale (1 L) preparations of BL21 (DE3) pGroESL and BL21-CodonPlus® (DE3) RP grown and induced as previously described. Cells were harvested and re-suspended in 75 mL of buffer A, and separated into three volumes of 25 mL. The suspensions were treated in three different ways, cell extract was lysed by ultrasonication (~50 % amplitude) at 4 ºC, either in 2 s pulses with 10 s delays, in 20 s pulses with 2 min delays or in 20 s pulses with 2 min delays in combination with BugBuster™ (2.5 mL of 1× stock) and lysozyme (0.4 mg). Two aliquots were taken from each treatment and strain combination after various lengths of sonication. One aliquot was centrifuged to pellet out insoluble protein. The pellet was re-suspended in an equivalent volume of 3 M urea and 10 % SDS. The total, soluble and insoluble components of each protein preparation were examined by SDS-PAGE.

6.5 Over-expression and purification of *E. coli* DHDPS

This protocol used was based on the procedure developed by Griffin *et al*.[1,36]

Buffer A	20 mM Tris.HCl, pH 8.0
Buffer B	20 mM Tris.HCl, pH 8.0 1 M NaCl
Buffer C	20 mM Tris.HCl, pH 8.0 0.5 M ammonium sulfate

6.5.1 Growth of *E. coli* XL1-Blue

Glycerol freezer stocks of *E. coli* XL1-Blue containing the appropriate plasmid were streaked out on an LB agar plate containing ampicillin and tetracycline and grown overnight at 37 ºC. Single colonies were selected to inoculate 3 mL LB broth, containing the same antibiotics, and starter cultures were incubated overnight at 37 ºC with shaking (200 rpm). Six flasks of 750 mL LB broth were supplemented with 750 µL of the appropriate antibiotics, inoculated with 1.0 mL of starter culture and grown at 37 ºC overnight with shaking.

The cells were chilled on ice for 30 min, harvested by centrifugation (5000 rpm, 15 min, 4 ºC), and washed with cold buffer A. The supernatant was discarded and the cell pellet, usually weighing ~10 g, was re-suspended in cold buffer A (1 mL per g of cells).

6.5.2 Preparation of crude extract

The cell suspension was ultrasonicated for 5 min at 4 °C in 2 s pulses with 10 s delays, using a Sonics Vibracell sonicator. Following sonication, the cell extract was clarified using centrifugation (10000 rpm, 15 min, 4 ºC), and an aliquot of 500 µL of the supernatant was reserved for analysis.

6.5.3 Heat shock

The supernatant was pipetted in 1 mL aliquots into sterile Eppendorf tubes. These were incubated for 2 min at 70 ºC and immediately cooled on ice. After centrifugation (14000 rpm,

10 min, 4 ºC), the supernatants were pooled, and a further 500 µL of the supernatant was reserved for analysis.

6.5.4 Ion exchange chromatography

The pooled sample from the heat shock step was loaded onto a Q-Sepharose ion exchange column (bed volume 75 mL, 15 × 2.6 cm) pre-equilibrated with buffer A, and subsequently washed with 5 column volumes of the same buffer. A salt gradient, increasing NaCl from 0 to 1 M, was applied with buffer B to elute proteins associating with column resin by ionic interactions. The peak fractions, as determined by absorbance at 280 nm, were tested for activity using the *o*-aminobenzaldehyde assay (described in section 6.4.4). Active fractions were pooled and 500 µL was set aside for further analysis, prior to dialyzing (detailed in section 6.4.5) the protein back into buffer A.

6.5.5 Hydrophobic interaction chromatography

Ammonium sulfate was added to pooled active fractions from ion exchange to give a final concentration of 0.5 M. This was loaded onto a phenyl Sepharose column (bed volume 125 mL, 25 × 2.6 cm) pre-equilibrated with buffer C, and subsequently washed with 5 column volumes of the same buffer. Proteins were eluted with a gradient decreasing in ammonium sulfate concentration from 0.5 to 0 M, applied by buffer A. The active fractions, as determined by the *o*-aminobenzaldehyde assay (described in section 6.4.4) were pooled and the protein was dialyzed (detailed in section 6.4.5) back into buffer A. Aliquots of 500 µL taken at different purification steps were analyzed by SDS-PAGE and by the Bradford and coupled assays.

6.6 Over-expression and purification of TEV protease

Purification of TEV protease involved a simplified protocol based on published methods.[16]

Buffer A	50 mM Na_2HPO_4 (dibasic phosphate), pH 8.0
	100 mM NaCl
	10 % (v/v) glycerol
	25 mM imidazole

Buffer B	50 mM Na_2HPO_4 (dibasic phosphate), pH 8.0 100 mM NaCl 10 % (v/v) glycerol 200 mM imidazole

6.6.1 Growth of *E. coli* BL21-CodonPlus (DE3) RIL

Starter cultures of *E. coli* BL21-CodonPlus (DE3) RIL pRK793 were prepared in an analogous way to that described in previous sections 6.4.1, using ampicillin and chloramphenicol as selective antibiotics. Four flasks containing 750 mL of LB broth were supplemented with 750 µL of ampicillin and chloramphenicol and inoculated with 1.5 mL of starter culture. These large cultures were grown at 37 ºC with shaking for ~10 hours.

Over-expression was induced with 750 µL of IPTG (60 $mg.mL^{-1}$, 0.25 M) after this initial growth period, and the cells were grown overnight at room temperature. The cells were then chilled on ice for 30 minutes, harvested by centrifugation (4000 rpm, 12 min, 4 ºC), and washed by re-suspension in cold buffer A, followed by further centrifugation (8000 rpm, 10 min, 4 ºC). The supernatant was discarded and the cell pellet, weighing ~6 g, was re-suspended in cold buffer A (4 mL per g of cells).

6.6.2 Preparation for affinity chromatography

The cell suspension was ultrasonicated for 5 min at 4 °C in 1.5 s pulses with 10 s delays, using a Sonics Vibracell sonicator. Following sonication, the cell extract was clarified using centrifugation (10000 rpm, 15 min, 4 ºC), and an aliquot of 500 µL of the supernatant was reserved for analysis.

6.6.3 Affinity chromatography

The supernatant was loaded on a His_6-tag affinity column (2× 5 mL HisTrap™ Crude or HisTrap™ FF, GE Healthcare) pre-equilibrated with buffer A, and subsequently washed with 5 column volumes of the same buffer. The His_6-tagged *T. maritima* DHDPR was eluted at high imidazole concentration by buffer B.

The peak fractions, as determined by absorbance at 280 nm, were analyzed by SDS-PAGE (section 6.3.2) for the presence of 27 kDa TEV protease, and buffer exchanged into buffer A. Protein concentration was quantified by the Bradford assay (section 6.3.1). The purified TEV protease was stabilized by the addition of EDTA and DTT to a concentration of 1 mM and stored at -80 ^{o}C.

6.7 Over-expression and purification of *T. maritima* DHDPR

Buffer A	50 mM Na_2HPO_4 (dibasic phosphate), pH 8.0 300 mM NaCl 30 mM imidazole
Buffer B	50 mM Na_2HPO_4 (dibasic phosphate), pH 8.0 300 mM NaCl 300 mM imidazole

Purification of *T. maritima* DHDPR was based on the procedure developed by Pearce *et al*.[11]

6.7.1 Growth of *E. coli* XL1-Blue and *E. coli* BL21 (DE3)

Large cultures of *E. coli* XL1-Blue pTM1520 or *E. coli* BL21 (DE3) pTOPO-TM1520 were prepared in an analogous way to that described in sections 6.5.1 and 6.4.1, respectively, using ampicillin and tetracycline as selective antibiotics. For 750 mL LB cultures of *E. coli* XL1-Blue pTM1520, over-expression was induced with 7.5 mL of arabinose (10 mg.mL^{-1}) after an initial growth period at 37 ºC and the cells were grown overnight at room temperature. For auto-inducing cultures of *E. coli* BL21 (DE3) pTOPO-TM1520, the flasks were moved to a room temperature shaker after an initial growth period and grown for an additional 12 hours. After growth the cultures were treated in the same manner. The cells were then chilled on ice for 30 min, harvested by centrifugation (4000 rpm, 15 min, 4 ºC), and washed with cold buffer A. The supernatant was discarded and the cell pellet, weighing ~20 g, was re-suspended in cold buffer A (2 mL per g of cells).

6.7.2 Preparation for affinity chromatography

The cell suspension was ultrasonicated in an analogous manner to that described in section 6.5.2. Following sonication, the cell extract was centrifuged (11000 rpm, 20 min, 4 ºC), and 500 µL of the supernatant was set aside for further analysis. Subsequently, the supernatant was incubated for 2 minutes at ~95 ºC, followed by centrifugation, with another 500 µL of the supernatant set aside for further analysis.

6.7.3 Affinity chromatography

The supernatant was loaded onto a His_6-tag affinity column (2× 5 mL HisTrap™ Crude or HisTrap™ FF, GE Healthcare) pre-equilibrated with buffer A, and subsequently washed with 5 column volumes of the same buffer. The His_6-tagged *T. maritima* DHDPR was eluted with high imidazole by buffer B**.**

The peak fractions, as determined by absorbance at 280 nm, were tested for activity using the coupled assay (described in section 6.9). Active fractions were pooled and the protein was dialyzed (detailed in section 6.4.5) into 20 mM Tris.HCl, pH 8.0. Aliquots of 500 µL taken at different steps were analyzed by SDS-PAGE, and by the Bradford and coupled assays.

6.8 Biophysical methods

6.8.1 Analytical gel-filtration liquid-chromatography

Analytical gel-filtration liquid-chromatography was performed using a pre-packed Superdex 200 (10/300 GL) column with a 24 mL bed volume and the UPC-900 ÄKTA FPLC from GE Healthcare at both 4 ºC and room temperature. The column was always pre-equilibrated with at least one column volume of 20 mM Tris.HCl, pH 8.0, containing varying concentrations of pyruvate, NaCl, glycerol and β-ME. The column was calibrated with BSA and ovalbumin, as they have been found to exhibit close to ideal elution behaviour on Sephadex columns.[37] For each run, 200 µL of protein sample, within the concentration range of ~1.5-0.15 $mg.mL^{-1}$, was loaded and eluted with the same buffer used to pre-equilibrate the column. Occasionally a preparative pre-packed Hi-load Superdex (16/60) column with a 120 mL bed volume was used at 4 ºC, which required calibration with BSA, ovalbumin and alcohol dehydrogenase.

The absorbance of the elutant was monitored at 280 nm or 280, 215, and 205 nm, resulting in chromatographs plotting absorbance as a function of buffer volume and showing the protein elution peaks. The peak of the elution curve was taken as the elution volume (V_e), which was used to either construct a calibration curve of V_e versus the natural log of the known molecular weight or to ascertain the molecular weight of unknown sample. Experimentally determined molar masses were compared to molar masses predicted from amino-acid sequences using the web-based program "Protein Calculator".[30]

6.8.2 Differential scanning fluorimetry

Based on the method developed by Ericsson *et al.*,[38] protein unfolding was monitored by fluorescence emission at 575 nm, using an iQ5 Real Time PCR Detection System (Bio-Rad). The thermal stabilities of the proteins of interest were assessed in 20 mM Tris.HCl, pH 8.0 solutions, containing varying concentrations of pyruvate, NaCl, glycerol and β-mercaptoethanol in a thin-walled 96 well PCR plate (Bio-Rad). Each 25 µL assay contained 1 µL of 250× SYPRO Orange fluorescent dye (Molecular Probes) to give a 10× final concentration. Each assay was performed in triplicate. Controls contained only buffer and fluorescent dye, or only buffer and protein. The plates were sealed with Microseal 'B' PCR sealer (Bio-Rad) and heated from 20 to 90 °C in increments of 0.2 ºC, with 20 s dwell times at each temperature step. Fluorescent changes were monitored simultaneously with a charge coupled device camera with an excitation filter of 490 nm and an emission filter of 575 nm. The melting temperature was determined as the point of maximum inflection of the fluorescence curve.[38]

6.8.3 Analytical ultracentrifugation

Analytical ultracentrifuge experiments were performed at the University of Melbourne in a Beckman Coulter XL-I analytical ultracentrifuge equipped with a UV/VIS absorbance optical system. The protein samples were desalted into 20 mM Tris.HCl, 150 mM NaCl, at pH 8.0 and loaded into 12 mm double sector cells with quartz windows, along with a reference solution of buffer without enzyme. The program SEDNTERP was used to determine the partial specific volume ($\bar{v}$) of the sample (wild-type: 0.7402 $mL.g^{-1}$, DHDPS-A204R: 0.7400 $mL.g^{-1}$ at 20 ºC), and the density (ρ) and viscosity (η) of the buffer (1.005 $g.mL^{-1}$ and 1.021 cP, respectively).[39]

Experiments were also performed with protein samples in storage buffer (20 mM Tris.HCl, 250 mM NaCl, 5 % (0.68 M) glycerol, 2 mM β-mercaptoethanol and 10 mM pyruvate at pH 8.0). The solution properties were estimated for storage buffer, ignoring the contributions by β-ME and pyruvate, because of the limitations of SEDNTERP. A density of 1.023 $g.mL^{-1}$ and viscosity of 1.193 cP were determined for a buffer composed of 20 mM Tris.HCl, 250 mM NaCl and 0.68 M glycerol using SEDNTERP.

Sedimentation velocity experiments

For sedimentation velocity (SV) experiments, protein sample (380 μL) and reference solution (400 μL) were loaded into double sector cells in an An-60 Ti 4-hole rotor and centrifuged at 40000 rpm. Continuous radial scans monitored sedimentation at 230 nm in low concentration samples (0.055 $mg.mL^{-1}$) and at 280 nm in high concentration samples (1.1 $mg.mL^{-1}$), every 8 min, without averaging, at 20 ºC. SV experiments were also performed in storage buffer, containing samples of both wild-type (1.1 $mg.mL^{-1}$) and A204R *M. tuberculosis* DHDPS (0.60 $mg.mL^{-1}$) samples and data were fitted to continuous size-distribution and continuous mass-distribution models using the program SEDFIT.[40]

In all cases, the standardised sedimentation coefficient ($s^{o}_{20,w}$) was determined using SEDNTERP.[39] Experimentally determined sedimentation coefficients were compared to sedimentation coefficients predicted from the PDB file 1XXX using hydrodynamic modelling with the program HYDROPRO.[41]

Sedimentation equilibrium experiments

For sedimentation equilibrium (SE) experiments, protein sample (100 μL) and reference solution (120 μL) were loaded into double sector cells in an An-60 Ti 8-hole rotor and centrifuged at 10000, 16000 or 23000 rpm until sedimentation equilibrium was attained (16-24 hrs). Radial absorbance scans were taken at 280 nm with 10 averages and 0.001 cm step size. Data were prepared for analysis with SEDFIT, and analyzed by globally fitting with various models using the program SEDPHAT.[42] Both SEDFIT and SEDPHAT are available from www.analyticalultracentrifugation.com.

6.8.4 Circular dichroism spectroscopy

Circular dichroism (CD) spectroscopy data were recorded at the University of Melbourne on an Aviv 60DS CD spectrophotometer at 20 ºC in 20 mM Tris.HCl, pH 8.0. The spectrophotometer was blanked with the buffer (20 mM Tris.HCl, pH 8.0). Wavelength scans were collected using a 1 mm path-length cuvette, 1.00 nm bandwidth, 0.5 nm step-size, from 250 to 200 nm with an averaging time of 2.0 s. Data were analyzed using CDPro software (freely available from http://lamar.colostate.edu/~sreeram/CDPro) to provide estimations of secondary structure fractions.[43,44] Three different algorithms, CONTINLL, SELCON3 and CDSSTR, were used to fit data points from 240 to 200 nm, in combination with the different reference sets provided by CDPro.[43,44]

6.9 Kinetic analysis of wild-type and mutant DHDPS

Kinetic parameters were determined using two-substrate steady-state initial-rate analysis. DHDPS activity was measured by coupling it to the oxidation of NADPH by DHDPR as previously described by Coulter *et al.*[8] The coupled assay used in this kinetic study was based on an optimized procedure for *E. coli* DHDPS, where the pH is maintained at 8.0 by 100 mM *N*-2-hydroxyethylpiperazine-*N*′-2-ethane sulfonic acid (HEPES) buffer.[45]

Assays were performed at a constant temperature of 30 ºC, using a circulating water bath, as described in the literature.[45] All components, except DHDPS and (*S*)-aspartate-β-semialdehyde, (*S*)-ASA, were incubated for 15 minutes at 30 ºC prior to initiation of the enzymatic reaction by a small aliquot of (*S*)-ASA. DHDPS was added just prior to (*S*)-ASA to minimize any possible degradation. (*S*)-ASA was synthesized using the methods outlined in Roberts *et al.*,[46] and was of high quality (>95 %) as judged by ^{1}H nuclear magnetic resonance spectroscopy (synthesized by Dr Sean Devenish and Dr Andy Muscroft-Taylor). Double distilled water was used as a blank and consumption of NADPH was monitored by change in absorbance at 340 nm. Control assays were performed routinely to ensure both the absence of contaminating NADPH-utilizing enzymes and that DHDPR was present in excess.

6.9.1 Enzyme stability and optimization

The optimum pH for enzyme activity was determined using a series of 20 mM buffers (MES, HEPES or Bicine) covering a pH range of 6 to 9 at 30 °C with ionic strength (IS) of 0.15 M

(adjusted by the addition of NaCl). The optimum IS for the enzymatic reaction was determined using a series of 20 mM HEPES buffers with varying NaCl concentrations, corrected for the amount of HCl/NaOH added to bring the pH to 8.25 at 30 °C.

For heat stability assays investigating wild-type DHDPS from *M. tuberculosis* and *E. coli*, enzymes were buffer exchanged into 20 mM HEPES (pH 8.25, IS of ~0.15) prior to incubation at various temperatures (30-100 °C) using a solid heat block. Aliquots of 10 µL were taken after 5 minutes incubation at various temperatures, or incubation at 70 °C for a range of times (0-30 minutes), and stored on ice or added directly to the coupled assay.

Heat stability assays comparing mutant A204R to wild-type *M. tuberculosis* DHDPS examined enzyme stored in 20 mM Tris.HCl, 250 mM NaCl, 5 % glycerol and 2 mM β-ME both with and without 10 mM pyruvate. The samples in the different buffers were pre-treated by 5 min incubation at a range of temperature (30-100 °C) and added directly to the coupled assay. These assays were performed at pH 8.0, buffered by 100 mM HEPES buffer.

6.9.2 Kinetic analysis for wild-type *M. tuberculosis* DHDPS

In contrast to most of the stability and optimization experiments, for kinetic analysis the coupled assay had an IS of ~0.15 M and a pH of 8.0, maintained by 100 mM HEPES buffer (from a stock of 200 mM HEPES, 50 mM NaCl). The coupled assay contained the following components: pyruvate, DHDPR, NADPH, HEPES buffer, (*S*)-ASA, and DHDPS. The nature of the assay allowed several components to be premixed. For a final cuvette volume of 1 mL, the addition of 580 µL of the premix gave the appropriate concentrations of HEPES, NADPH and DHDPR, as outlined in **Table 6.11**.

Table 6.11: Preparation for kinetic analysis of *M. tuberculosis* DHDPS. The premix stock is sufficient for 95 individual assays.

	[Stock] mM	Per cuvette (µL)	[Cuvette] mM	Premix (mL)
HEPES, pH 8.0 (and 50 mM NaCl)	200	500	100	47.5
NADPH	6.0	30	0.18	2.85
DHDPR (0.13 mg.mL^{-1})	--	50	--	4.75
Total	--	580	--	55.1

Since the activity of DHDPS was monitored by coupling it to the consumption of NADPH by DHDPR, care was taken to ensure an excess of the coupling enzyme. Purified *T. maritima* DHDPR was added to the premix to give DHDPR activity in the cuvette which was at least four times greater than the largest DHDPS activity to be measured. To confirm this excess, the addition of more DHDPR was found to not alter measured rates. Controls were also performed to ensure the absence of contaminating NADPH-utilizing enzymes.

To determine the substrate concentrations to be used for accurate determination of K_M, apparent K_M values were determined by varying one substrate while holding the other constant, using as wide a range of substrate concentrations as practically possible. Determination of true kinetic constants, K_M and V, involved concurrent variation of both substrate concentrations using 0.2 to 10 times that of the apparent K_M values. For wild-type DHDPS, pyruvate and (*S*)-ASA stocks were both added to the cuvette to give final concentrations covering the range from 0.060 to 3.00 mM (that is 0.060, 0.15, 0.30, 1.5 and 3.0 mM).

Preliminary assays were performed in which (*S*)-ASA was the limiting factor in order to accurately determine the concentration of (*S*)-ASA from the total amount of NADPH consumed. The concentration of DHDPS was determined by Bradford assays (section 6.3.1), and it was shown to be of high purity with SDS-PAGE (section 6.3.2). Wild-type *M. tuberculosis* DHDPS was added just prior to the addition of (*S*)-ASA, giving a protein concentration of 0.0011 $mg.mL^{-1}$ in each assay cuvette.

A total of 25 different substrate concentrations were assayed. The initial-rate of NADPH consumption was measured in at least triplicate and was reproducible (±10 %). The data were fitted with kinetic models using the program ENZFITTER (Biosoft, Cambridge, U.K.)

6.9.3 Inhibition studies with (*S*)-lysine and other amino acids

The effect of (*S*)-lysine on the activity of *M. tuberculosis* DHDPS was examined using the coupled assay. The assay was performed in an analogous way to that previously described, but with the addition of (*S*)-lysine at concentrations ranging from 0 to 400 mM. The concentrations of the substrates (0.32 mM pyruvate, 0.20 mM (*S*)-ASA) were kept constant and the activities for each (*S*)-lysine concentration were determined in triplicate and reported

relative to that observed in the absence of the inhibitor. The IC_{50}, that is the concentration needed to inhibit enzyme activity by 50 %, was determined by plotting relative activity against (*S*)-lysine concentration.

To investigate the effect of other aspartate family amino acids on *M. tuberculosis* DHDPS activity, the coupled assay was performed, with the addition of either meso-DAP, (*S*)-threonine, or (*S*)-methionine at concentrations of 1, 10, 20 mM or in combinations of two for (*S*)-lysine, meso-DAP, (*S*)-threonine, or (*S*)-methionine at concentrations of 5 mM.. The concentrations of the substrates (0.30 mM pyruvate, 0.18 mM (*S*)-ASA) were kept constant and the activities for each amino acid concentration were determined in triplicate/duplicate and reported relative to that observed in the absence of the inhibitor.

6.9.4 Kinetic analysis for mutant *M. tuberculosis* DHDPS

Kinetic parameters were also determined for mutant A204R *M. tuberculosis* DHDPS, and the coupled assays were performed as described in section 6.9.2, with a few distinctions. DHDPS-A204R was stored in buffer containing 10 mM pyruvate and therefore the 5 µL aliquot of enzyme added just prior to (*S*)-ASA, increased pyruvate concentration by 0.05 mM. This addition was taken into account when preparing stock solutions, and the final pyruvate concentrations in the cuvette were 0.50, 0.10, 0.20, 1.0, and 2.0 mM. These concentrations were chosen based on apparent K_M values determined for DHDPS-A204R. The high apparent K_M value determined for (*S*)-ASA meant the cuvette concentrations of (*S*)-ASA ranged from 0.22 to 5.50 mM; thus a higher concentration of HEPES was required to buffer pH at 8.0, giving an ionic strength of 0.18 M, rather than 0.15 M. Consequently, the premix composed of HEPES, NADPH, and DHDPR was prepared as outlined in **Table 6.12**.

Table 6.12: Preparation for kinetic analysis of A204R *M. tuberculosis* DHDPS, providing enough premixed components for 100 individual assays.

	[Stock] mM	Per cuvette (µL)	[Cuvette] mM	Premix (mL)
HEPES, pH 8.0 (and 50 mM NaCl)	200	750	150	75.0
NADPH	6.0	25	0.15	2.5
DHDPR (0.58 $mg.mL^{-1}$)	--	40	--	4.0
Total	--	815	--	81.5

As in kinetic studies with wild-type *M. tuberculosis* DHDPS, all components except DHDPS and (*S*)-ASA were incubated for 15 minutes at 30 °C. Prior to the addition of (*S*)-ASA, which initiated the reaction, A204R *M. tuberculosis* DHDPS was added, giving a protein concentration of 0.0026 $mg.mL^{-1}$ in each assay cuvette.

6.10 X-ray crystallography

X-ray crystallography of the mutant *M. tuberculosis* DHDPS-A204R was performed in collaboration with Linda Schuldt, under the supervision of Dr Manfred Weiss, from the EMBL Hamburg Outstation. The protein was purified in our lab (as described in section 6.4) and crystallization experiments were performed by our collaborators at the EMBL High Throughput Crystallization Facility, in Hamburg, Germany **(**details in **Appendix I)**.[47]

Diffraction data were collected by our collaborators (as described in section 6.10.1), and the structure was solved by molecular replacement using the program MOLREP,[48] with a search model consisting of the chains *a* and *b* from the crystal structure of wild-type *M. tuberculosis* DHDPS (PDB entry 1XXX). Initial refinement was performed by our collaborators, using the program REFMAC5.[48,49] Subsequent refinement was performed by myself, under the supervision of Prof. Geoff Jameson and Dr Sean Devenish (section 6.10.2).

6.10.1 Diffraction data collection and processing

Compact crystals were grown in 2.0 M ammonium sulfate and 100 mM sodium acetate, pH 5.5 at room temperature within 4 weeks to a maximum size of 100 × 40 × 20 μm^3. Diffraction data were collected by our collaborators using a single crystal on a beamline equipped with a MARMosaic CCD-detector. Data collection was performed as outlined in **Table 6.13**. The crystal was mounted in a cryo-loop, soaked quickly in 20 % (v/v) glycerol in reservoir solution and flash cooled in a nitrogen stream to 100 K. A total of 221 images were collected with a rotation increment of 0.5° and an exposure time of 30 seconds per image. Due to radiation damage, however, only the first 120 images (60° of data) were used for final data processing. Indexing and integration of the data were done using the program DENZO, followed by scaling with SCALEPACK.[50]

Table 6.13: Data collection for DHDPS-A204R.

Date	13.06.2008
No. of crystals	1
Beamline	I911-3 (Max Lab, Lund, Sweden)
Detector	Mar Mosaic CCD 225 mm
Wavelength (Å)	1.000
Crystal-detector distance (mm)	225

Intensities were converted to structure factor amplitudes using the program TRUNCATE.[48] The optical resolution of the data set was calculated with SFCHECK,[51] and the self rotation function was computed using the program MOLREP,[48] based on structure factor amplitudes to a maximum resolution of 4.0 Å. The *R*-factors $R_{r.i.m.}$ (redundancy-independent merging *R*-factor) as well as $R_{p.i.m.}$ (precision-indicating merging *R*-factor)[52] were calculated using the program RMERGE (available from http://www.embl-hamburg.de/~msweiss/projects/msw_qual.html or from Dr Manfred Weiss upon request) and are listed in **Table 6.14**.

Table 6.14: *R*-factors for DHDPS-A204R. Values in parentheses are for the highest resolution bin.

R_{merge} (%)	11.8 (58.9)
$R_{r.i.m.}$ (%)	13.3 (65.9)
$R_{p.i.m.}$ (%)	5.9 (29.7)

6.10.2 Structural refinement

Partially refined model was received from Germany with *R*- and free *R*-factors of 23.6 % and 28.4 %, respectively. Further refinement was achieved using REFMAC5 with applied NCS restraints and anisotropic displacement parameters, using the TLS (translation, rotation, screw-rotation) model.[48,53] Manual model corrections were made using the program WinCOOT.[54] The final rounds of refinement involved the placement of solvent molecules and waters using WINCOOT. Structural quality was assessed using SFCHECK,[51] and the structural validation tools of WINCOOT. In some cases, the side chains of amino acids could not be completely resolved due to insufficient electron density so the occupancy of these atoms was lowered. The maximum r.m.s.d. for the subunits (all atoms) in the asymmetric unit was 0.38 Å, as calculated by SUPERPOSE.[48]

6.11 References

1. Griffin, M. D. W. Why is dihydrodipicolinate synthase a tetramer? *D. Phil Thesis*, Univeristy of Canterbury. (2005).

2. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular cloning: a laboratory manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989).

3. Goloubinoff, P., Gatenby, A. A. & Lorimer, G. H. GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose bisphosphate carboxylase oligomers in *Escherichia coli*. *Nature* **337**, 44-7 (1989).

4. Kraunsoe, J. A. E. Studies in lysine biosynthesis. *BSc. Hons. Part II*, Corpus Christi College, Oxford. (1992).

5. Gerrard, J. A. Studies on dihydrodipicolinate synthase. *D. Phil Thesis*, Oxford. (1992).

6. Coulter, C. V. Studies in (*S*)-lysine biosynthesis via the diaminopimelate pathway. *D. Phil Thesis*, Canterbury. (1997).

7. Coulter, C. V., Gerrard, J. A., Kraunsoe, J. A. E., Moore, D. J. & Pratt, A. J. *(S)*-Aspartate semi-aldehyde: Synthetic and structural studies. *Tetrahedron* **52**, 7127-36 (1996).

8. Coulter, C. V., Gerrard, J. A., Kraunsoe, J. A. E. & Pratt, A. J. *Escherichia coli* dihydrodipicolinate synthase and dihydrodipicolinate reductase: kinetic and inhibition studies of two putative herbicide targets. *Pesticide science* **55**, 887-95 (1999).

9. Dubendorff, J. W. & Studier, F. W. Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with *lac* repressor. *Journal of molecular biology* **219**, 45-59 (1991).

10. Novagen. *pET system tutorial manual, 11th edition* (2006).

11. Pearce, F. G., Sprissler, C. & Gerrard, J. A. Characterization of dihydrodipicolinate reductase from *Thermotoga maritima* reveals evolution of substrate binding kinetics. *Journal of biochemistry* **143**, 617-23 (2008).

12. Lesley, S. A., Kuhn, P., Godzik, A., Deacon, A. M., Mathews, I., Kreusch, A., Spraggon, G., Klock, H. E., McMullan, D., Shin, T., Vincent, J., Robb, A., Brinen, L. S., Miller, M. D., McPhillips, T. M., Miller, M. A., Scheibe, D., Canaves, J. M., Guda, C., Jaroszewski, L., Selby, T. L., Elsliger, M. A., Wooley, J., Taylor, S. S., Hodgson, K. O., Wilson, I. A., Schultz, P. G. & Stevens, R. C. Structural genomics of the *Thermotoga maritima* proteome implemented in a high-throughput structure determination pipeline. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 11664-9 (2002).

13. Kefala, G. & Weiss, M. S. Cloning, expression, purification, crystallization and preliminary X-ray diffraction analysis of *dap*A (Rv2753c) from *Mycobacterium tuberculosis*. *Acta crystallographica. Section F, Structural biology and crystallization communications* **62**, 1116-9 (2006).

14. Dummler, A., Lawrence, A. M. & de Marco, A. Simplified screening for the detection of soluble fusion constructs expressed in *E. coli* using a modular set of vectors. *Microbial cell factories* **4**, 34 (2005).

15. Chang, A. C. & Cohen, S. N. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *Journal of bacteriology* **134**, 1141-56 (1978).

16. Kapust, R. B., Tozser, J., Fox, J. D., Anderson, D. E., Cherry, S., Copeland, T. D. & Waugh, D. S. Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein engineering* **14**, 993-1000 (2001).

17. Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein expression and purification* **41**, 207-34 (2005).

18. Stratagen. Quikchange site-directed mutagenesis kit: instruction manual. (2004).

19. Sambrook, J., Fritsch, E. F. & Maniatis, T. Small-scale preparation of plasmid DNA in *Molecular cloning: a laboratory manual* (ed. Nolan, C.) 1.25-1.8 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989).

20. QIAGEN. QIAquick® Spin Handbook. (2002).

21. Sambrook, J., Fritsch, E. F. & Maniatis, T. Preparation of reagents and buffers used in molecular cloning in *Molecular cloning: a laboratory manual* (ed. Nolan, C.) B.1-B.28 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989).

22. Sanger, F., Nicklen, S. & Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **74**, 5463-7 (1977).

23. Deutscher, M. P. *Guide to protein purification* (Academic Press limited, Sydney, 1990).

24. Scopes, R. K. *Protein purification: principles and practice* (Springer-Verlag, New York, 1987).

25. Scopes, R. K. *Protein purification: principles and practice* (Springer-Verlag, New York, 1994).

26. Eisenthal, R. & Danson, M. J. *Enzyme assays: a practical approach* (IRL Press at Oxford University Press, Oxford [England]; New York, 1992).

27. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry* **72**, 248-54 (1976).

28. Stoscheck, C. M. Quantitation of Protein in *Guide to protein purification* (ed. Deutscher, M. P.) 50-68 (Academic Press, Sydney, 1990).

29. Zor, T. & Selinger, Z. Linearization of the Bradford protein assay increases its sensitivity: theoretical and experimental studies. *Analytical biochemistry* **236**, 302-8 (1996).

30. Putnam, C. D. Protein Calculator v3.3. in *http://www.scripps.edu/~cdputnam/protcalc.html* (The Scripps Research Institute, 2006).

31. Garfin, D. E. One-dimensional gel electrophoresis in *Guide to protein purification* (ed. Deutscher, M. P.) 425-41 (Academic Press, Sydney, 1990).

32. Invitrogen. Catalogue 2004 Invitrogen: Solutions for Life Science Research and Drug Discovery. (2004).

33. Schagger, H. & von Jagow, G. Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Analytical biochemistry* **199**, 223-31 (1991).

34. Yugari, Y. & Gilvarg, C. The condensation step in diaminopimelate synthesis. *The Journal of biological chemistry* **240**, 4710-6 (1965).

35. Invitrogen. *AcTEV™ protease: instruction manual* (2005).

36. Griffin, M. D., Dobson, R. C., Pearce, F. G., Antonio, L., Whitten, A. E., Liew, C. K., Mackay, J. P., Trewhella, J., Jameson, G. B., Perugini, M. A. & Gerrard, J. A. Evolution of quaternary structure in a homotetrameric enzyme. *Journal of molecular biology* **380**, 691-703 (2008).

37. Andrews, P. Estimation of the molecular weights of proteins by Sephadex gel-filtration. *The Biochemical journal* **91**, 222-33 (1964).

38. Ericsson, U. B., Hallberg, B. M., Detitta, G. T., Dekker, N. & Nordlund, P. Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Analytical biochemistry* **357**, 289-98 (2006).

39. Laue, T. M., Shah, D. B., Ridgeway, T. M. & Pelletier, S. L. *Analytical ultracentrifugation in biochemistry and protein science* (The Royal Society of Chemistry, Cambridge, 1992).

40. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophysical journal* **78**, 1606-19 (2000).

41. Garcia De La Torre, J., Huertas, M. L. & Carrasco, B. Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophysical journal* **78**, 719-30 (2000).

42. Vistica, J., Dam, J., Balbo, A., Yikilmaz, E., Mariuzza, R. A., Rouault, T. A. & Schuck, P. Sedimentation equilibrium analysis of protein interactions with global implicit mass conservation constraints and systematic noise decomposition. *Analytical biochemistry* **326**, 234-56 (2004).

43. Sreerama, N., Venyaminov, S. Y. & Woody, R. W. Analysis of protein circular dichroism spectra based on the tertiary structure classification. *Analytical biochemistry* **299**, 271-4 (2001).

44. Sreerama, N. & Woody, R. W. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Analytical biochemistry* **287**, 252-60 (2000).

45. Dobson, R. C., Gerrard, J. A. & Pearce, F. G. Dihydrodipicolinate synthase is not inhibited by its substrate, (*S*)-aspartate β-semialdehyde. *The Biochemical journal* **377**, 757-62 (2004).

46. Roberts, S. J., Morris, J. C., Dobson, R. C. & Gerrard, J. A. The preparation of *(S)*-aspartate semi-aldehyde appropriate for use in biochemical studies. *Bioorganic & medicinal chemistry letters* **13**, 265-7 (2003).

47. Mueller-Dieckmann, J. The open-access high-throughput crystallization facility at EMBL Hamburg. *Acta crystallographica. Section D, Biological crystallography* **62**, 1446-52 (2006).

48. The CCP4 suite: programs for protein crystallography. *Acta crystallographica. Section D, Biological crystallography* **50**, 760-3 (1994).

49. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta crystallographica. Section D, Biological crystallography* **53**, 240-55 (1997).

50. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods in enzymology* **276**, 307-26 (1997).

51. Vaguine, A. A., Richelle, J. & Wodak, S. J. SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta crystallographica. Section D, Biological crystallography* **55**, 191-205 (1999).

52. Weiss, M. S. *Journal of applied crystallography* **34**, 130-5 (2001).

53. Winn, M. D., Murshudov, G. N. & Papiz, M. Z. Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods in enzymology* **374**, 300-21 (2003).

54. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta crystallographica. Section D, Biological crystallography* **60**, 2126-32 (2004).

Appendix A

The pET (inducible) protein expression system

Recombinant DNA technology has resulted in great advances in protein purification including inducible protein expression systems.[1] The level of gene transcription, and consequently protein production, depends on the frequency with which RNA polymerase initiates transcription, which is a function of the promoter sequence upstream from the gene.[1] The *dap*A gene **(Figure A.2)** for DHDPS from *M. tuberculosis* strain H37Rv was cloned into the expression vector pETM-11 (**Figure A.1**), by our collaborators, Weiss and Kefala.[2] The pETM vectors are derived from Novagen's pET expression system,[3] which uses a strong promoter from phage T7 recognized by T7 RNA polymerase.[4]

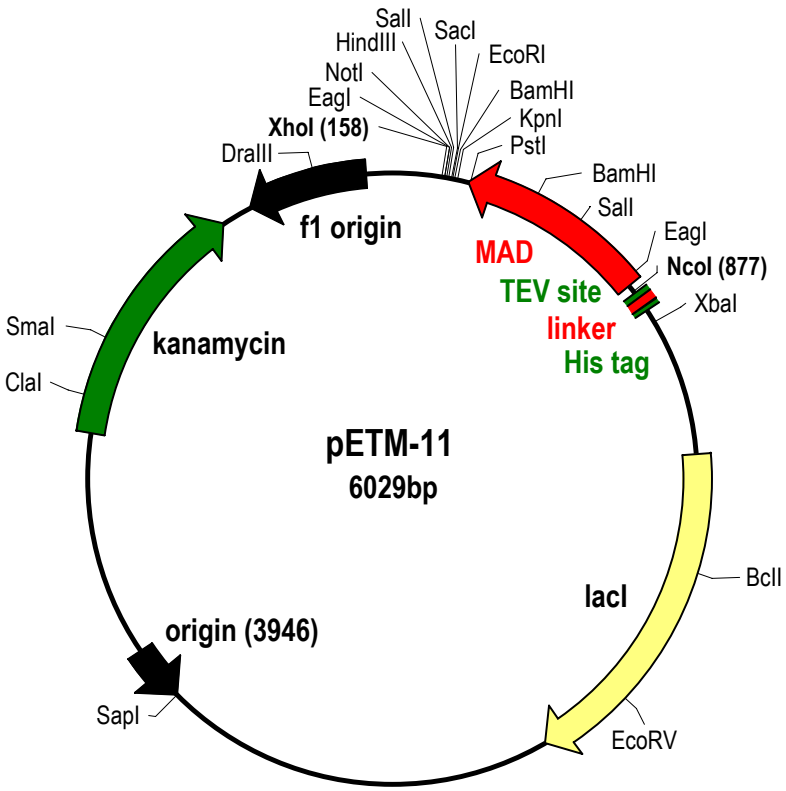


Figure A.1: Map of the pETM-11 vector developed by the EMBL Protein Expression and Purification Facility.[3] It is modified from pET24d (Novagen) with a 612 bp MAD insert, which contains additional restriction sites. The *dap*A gene from *M. tuberculosis* was cloned into the vector using the bolded restriction sites, *Nco*I and *Xho*I. The full sequence is available online at www.embl-hamburg.de/~geerlof/webPP/vectordb/bact_vectors/index.html.

```
1      GTGACCACCG TCGGATTCGA CGTCGCAGCG CGCCTAGGAA CCCTGCTGAC CGCGATGGTG
61     ACACCGTTTA GCGGCGATGG CTCCCTGGAC ACCGCCACCG CGGCGCGGCT GGCCAACCAC
121    CTGGTCGATC AGGGGTGCGA CGGTCTGGTG GTCTCGGGCA CCACCGGCGA GTCGCCGACC
181    ACCACCGACG GGGAGAAAAT CGAGCTGCTG CGGGCCGTCT TGGAAGCGGT GGGGGACCGG
241    GCCCGTGTTA TCGCCGGTGC CGGCACCTAT GACACCGCGC ACAGCATCCG GCTGGCCAAG
301    GCTTGTGCGG CCGAGGGTGC GCACGGGCTG CTGGTGGTCA CGCCCTACTA TTCCAAGCCG
361    CCGCAGCGGG GGCTGCAAGC CCATTTCACC GCCGTCGCCG ACGCGACCGA GCTGCCGATG
421    CTGCTCTATG ACATCCCGGG GCGGTCGGCG GTGCCGATCG AGCCCGACAC GATCCGCGCG
481    TTGGCGTCGC ATCCGAACAT CGTCGGAGTC AAGGACGCCA AAGCCGACCT GCACAGCGGC
541    GCCCAAATCA TGGCCGACAC CGGACTGGCC TACTATTCCG GCGACGACGC GCTCAACCTG
601    CCCTGGCTGG CCATGGGCGC CACGGGCTTC ATCAGCGTGA TTGCCCACCT GGCAGCCGGG
661    CAGCTTCGAG AGTTGTTGTC CGCCTTCGGT TCTGGGGATA TCGCCACCGC CCGCAAGATC
721    AACATTGCGG TCGCCCCGCT GTGCAACGCG ATGAGCCGCC TGGGTGGGGT GACGTTGTCC
781    AAGGCGGGCT TGCGGCTGCA GGGCATCGAC GTCGGTGATC CCCGGCTGCC CCAGGTGGCC
841    GCGACACCGG AGCAGATCGA CGCGTTGGCC GCCGACATGC GCGCGGCCTC GGTGCTTCGG
```

Figure A.2: The *dap*A gene as cloned into the pETM-11 vector system, and determined by DNA sequencing (as described in section 6.2.13).

T7 RNA polymerase does not naturally occur in *E. coli*, so gene products in pETM or pET vectors can only be expressed in bacterial strains engineered to contain its gene,[5] such as *E. coli* BL21(DE3).[6] The transcription of both the gene for T7 RNA polymerase and the cloned gene are blocked by the binding of a *lac* repressor to a *lac* operator element in their promoters.[6,7] Hence, suppressing the binding of the *lac* repressor, by its natural substrate allolactose or an artificial mimic, isopropyl thiogalactoside (IPTG), induces the expression of T7 RNA polymerase and the cloned gene product, typically a recombinant protein.[7] T7 RNA polymerase is so active and selective that, after being induced, almost all of the host cell's resources are used for recombinant protein expression and, after a few hours, the cloned gene product comprises up to 50 % of the total cellular proteins.[7]

Inducible protein expression allows some degree of control over recombinant protein production, and prevents cellular resources being diverted while cells are still dividing, during the lag phase of culture growth. Typically, IPTG was added after and OD_{600} of ~0.60 was reached. Auto-induction media take advantage of allolactose being a minor by-product of lactose metabolism,[8] and uses glucose to suppress its production during lag phase growth allowing for the growth of a high-density culture before induction of protein expression.[9]

References

1. Das, A. Overproduction of proteins in *Escherichia coli*: vectors, hosts, and strategies in *Guide to protein purification* (ed. Deutscher, M. P.) 93-112 (Academic Press, Sydney, 1990).

2. Kefala, G. & Weiss, M. S. Cloning, expression, purification, crystallization and preliminary X-ray diffraction analysis of *dap*A (Rv2753c) from *Mycobacterium tuberculosis*. *Acta crystallographica. Section F, Structural biology and crystallization communications* **62**, 1116-9 (2006).

3. Dummler, A., Lawrence, A. M. & de Marco, A. Simplified screening for the detection of soluble fusion constructs expressed in *E. coli* using a modular set of vectors. *Microbial cell factories* **4**, 34 (2005).

4. Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods in enzymology* **185**, 60-89 (1990).

5. Studier, F. W. & Moffatt, B. A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of molecular biology* **189**, 113-30 (1986).

6. Dubendorff, J. W. & Studier, F. W. Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with *lac* repressor. *Journal of molecular biology* **219**, 45-59 (1991).

7. Novagen. *pET system tutorial manual, 11th edition* (2006).

8. Mathews, C. K., Van Holde, K. E. & Ahern, K. G. *Biochemistry* (Benjamin/Cummings Pub. Co., Sydney, 2000).

9. Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein expression and purification* **41**, 207-34 (2005).

Appendix B

Kinetic characterization of His_6-tagged DHDPS

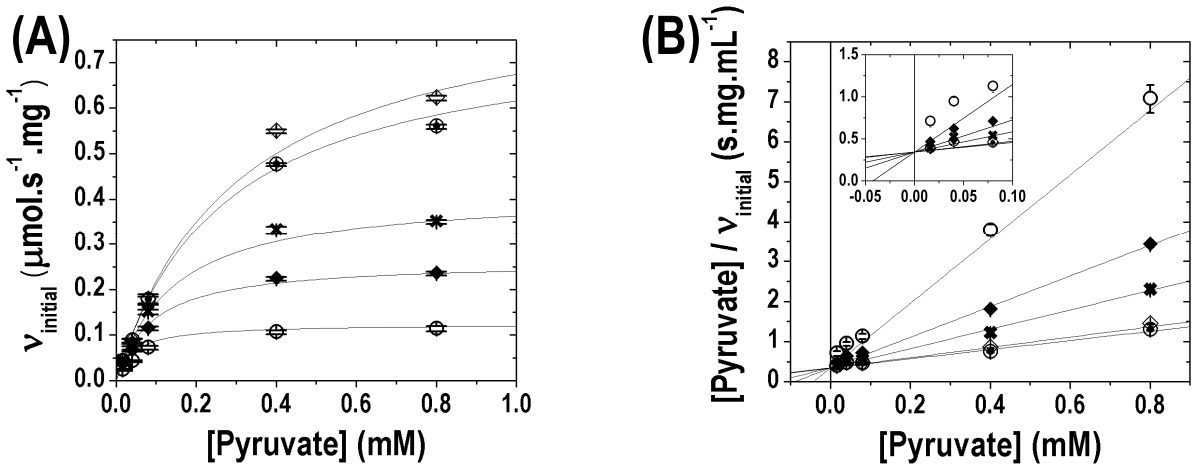


Figure B.1: **The initial velocity of His_6-tagged *M. tuberculosis* DHDPS was measured with the coupled assay over varying concentrations of both substrates, pyruvate and (*S*)-ASA, shown on the x-axis and with symbols (◇ 2.0 mM, ◉ 1.0 mM, × 0.20 mM, ◆ 0.10 mM, ○ 0.04 mM (*S*)-ASA). Each measurement was made in triplicate and the data were fitted with the ping-pong model giving an R^2 of 0.99 using ENZFITTER. Plot (B) shows the Hanes transformation, reflecting the trend of y-axis intersection predicted for the ping-pong model.**

Table B.1: **Kinetic parameters determined by fitting rate measurements.**

	His_6-tagged *M. tuberculosis* DHDPS	Non-tagged *M. tuberculosis* DHDPS
Kinetic model	Ping-pong	Ping-pong
V	1.00 (±0.04)	4.2 (±0.1)
K_M for (*S*)-ASA (mM)	0.28 (±0.02)	0.43 (±0.02)
K_M for pyruvate (mM)	0.35 (±0.01)	0.17 (±0.01)
Monomer molecular weight (Da)	34173.8	31156.5
e_0 (mgs of active site per µmol)	34.17	31.16
k_{cat} per active site (s^{-1})	34 (±1)	138 (±2)
k_{cat}/K_M for (*S*)-ASA ($s^{-1}.mM^{-1}$)	122 (±4)	320 (±20)
k_{cat}/K_M for pyruvate ($s^{-1}.mM^{-1}$)	98 (±4)	820 (±60)

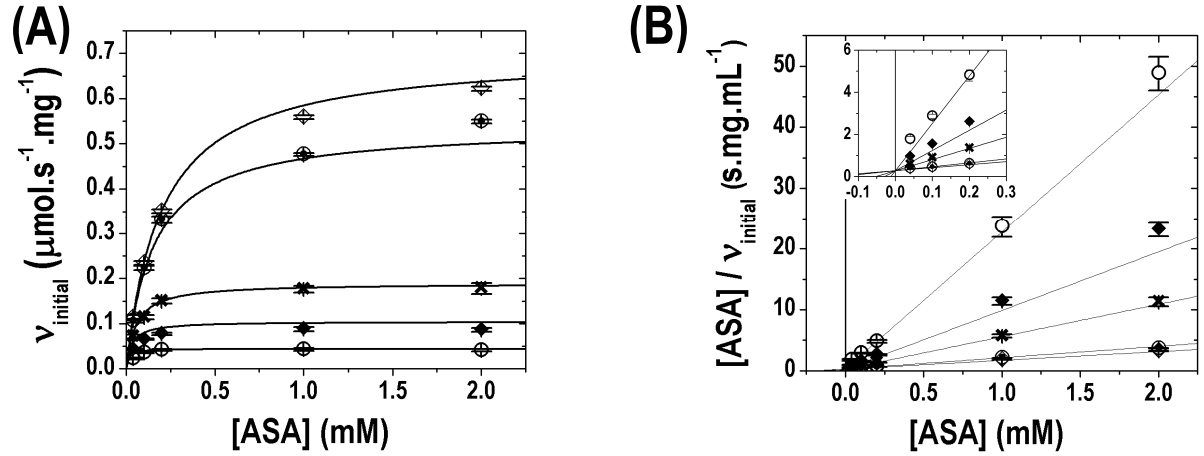


Figure B.2: **The initial velocity of His_6-tagged *M. tuberculosis* DHDPS was measured with the coupled assay over varying concentrations of both substrates, (*S*)-ASA and pyruvate, shown on the x-axis and with symbols (◇ 0.8 mM, ◉ 0.4 mM, ✕ 0.08 mM, ◆ 0.04 mM, ○ 0.016 mM pyruvate). Each measurement was made in triplicate and the data were fitted with the ping-pong model giving an R^2 of 0.99 using ENZFITTER. Plot (B) shows the Hanes transformation, reflecting the trend of y-axis intersection predicted for the ping-pong model.**

Appendix C

His_6-tag cleavage with TEV protease

In the vector pETM-11 **(Appendix A)**, *M. tuberculosis* DHDPS is expressed as a fusion protein, with a His_6-tag at the N-terminus linked using the sequence E-N-L-Y-F-Q-G-A-M-A. The nuclear inclusion protease from tobacco TEV has stringent sequence specificity, cleaving between Q and G residues in the sequence E-N-L-Y-F-Q-G, and consequently TEV protease is widely used for the removal of protein tags.[1] Incubation of TEV protease with the fusion protein removes the His_6-tag and most of the linker sequence; however, four linker amino acids, G-A-M-A, remain at the N-terminus of *M. tuberculosis* DHDPS.

The advantage of the His_6-tag is that it binds Ni^{2+} cations that are immobilized on chromatography resin with high specificity, giving a rapid one-step method of achieving greatly increasing protein purity.[2] The X-ray crystallography structure determined by Weiss and Kefala, shows both termini of *M. tuberculosis* DHDPS on the outer surface of the tetrameric structure. Therefore, the His_6-tag would not be expected to interfere with enzyme folding, but somewhat surprisingly a noticeable effect on kinetics was observed **(Appendix B)**.

Initially, cleavage was performed using commercially available TEV protease, called AcTEV™ protease, from Invitrogen, which is more stable than the native protease,[3] but susceptible to self-cleavage.[4] Kapust and colleagues rationally designed a mutant TEV-protease, S219V, which is 100-fold more stable than wild-type protease and does not self-cleave.[4] The plasmid pRK793 expresses the mutant TEV protease with a His_6-tag on the N-terminus, for purification and solubility purposes,[4] and was acquired through the non-profit organization Addgene (http://www.addgene.org/pgvec1). TEV protease was expressed and purified from BL21-CodonPlus® (DE3)-RIL cells containing pRK793, using a simplified protocol (detailed in chapter six, section 6.6) based on published methods,[4] and stored at -80 ^{o}C. Over-expression of TEV protease was apparent in the crude extract produced by sonication, as shown by SDS-PAGE with Coomassie blue staining **(Figure C.1, lane 2)**. The absence of the TEV protease in the flow-through from the His_6-tag affinity column **(Figure C.1, lane 4)** showed that all protein binds, and subsequent elution by high

concentrations (200 mM) of imidazole resulted in homogeneous TEV protease, as judged by SDS-PAGE **(Figure C.1, lane 5)**.

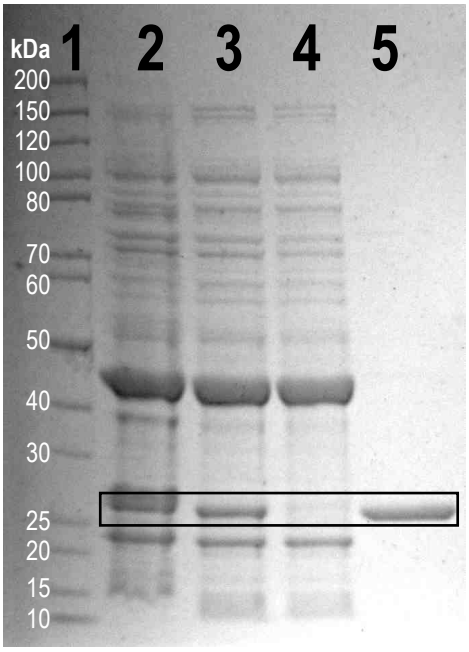


Figure C.1: The purification of recombinant TEV protease examined using denaturing SDS-PAGE, which separates proteins based on sub-unit molecular weight (MW). TEV protease was identified as the 27 kDa band (black rectangle) using the MW marker in lane 1. Protein preparations from various stages of purification were loaded into other lanes as follows: 2 - crude cell-free extract, 3 - after centrifugation, 4 - flow through from His_6-tag affinity column, 5 - eluted from His_6-tag affinity column.

References

1. Parks, T. D., Howard, E. D., Wolpert, T. J., Arp, D. J. & Dougherty, W. G. Expression and purification of a recombinant tobacco etch virus NIa proteinase: biochemical analyses of the full-length and a naturally occurring truncated proteinase form. *Virology* **210**, 194-201 (1995).

2. GE_Healthcare. Recombinant protein purification handbook: principles and methods. (2000).

3. Nayak, S., Li, L. & Lee, J. Enhanced TEV protease extends enzyme stability for long-term activity. *Focus* **25**, 12-4 (2003).

4. Kapust, R. B., Tozser, J., Fox, J. D., Anderson, D. E., Cherry, S., Copeland, T. D. & Waugh, D. S. Tobacco etch virus protease: mechanism of autolysis and rational design of stable mutants with wild-type catalytic proficiency. *Protein engineering* **14**, 993-1000 (2001).

Appendix D

Ionic strength in the coupled assay

The ionic strength of 200 mM HEPES, pH 8.0, was calculated using the buffer calculator (from www.liv.ac.uk/buffers), taking into consideration that the buffer was prepared at 20 °C, but the assay is performed at 30 °C.

Table D.1: Assay setup for wild-type *M. tuberculosis* DHDPS

Stock	[stock]		Volume µL	[cuvette]	
	mM	mg.mL^{-1}		mM	mg.mL^{-1}
HEPES, pH 8	200	-	500	100	-
Pyruvate	*varied*	-	*varied*	*varied*	-
NADPH	6.0	-	30	0.18	-
DHDPR	-	0.3	30	-	0.0040
d-H_2O	-	-	*varied*	-	-
(*S*)-ASA	*varied*	-	*varied*	*varied*	-
DHDPS	-	-	10	-	-
Total	-	-	**1000**	-	-

Table D.2: Assay setup for *M. tuberculosis* DHDPS-A204R

Stock	[stock]		Volume µL	[cuvette]	
	mM	mg.mL^{-1}		mM	mg.mL^{-1}
HEPES, pH 8	200	-	750	150	-
Pyruvate	*varied*	-	*varied*	*varied*	-
NADPH	6.0	-	30	0.18	-
DHDPR	-	0.13	30	-	0.0040
d-H_2O	-	-	*varied*	-	-
[(*S*)-ASA]	*varied*	-	*varied*	*varied*	-
DHDPS	-	-	10	-	-
Total	-	-	**1000**	-	-

Table D.3: Ionic strength (IS) in the cuvette

	Buffer stock components:		IS (M) added to buffer stock as:			Overall IS (M) of buffer stock	Overall IS (M) of assay mixture
	[HEPES] (mM)	[NaCl] (mM)	HEPES[a]	NaCl[b]	pH adjustment[b]		
Wild-type DHDPS	200	50	0.156	0.025	0.060	0.241	0.121
DHDPS-A204R	200	50	0.156	0.025	0.060	0.241	0.181

[a] Ionic strength calculated using the buffer calculator (from www.liv.ac.uk/buffers).

[b] Ionic strength $= \frac{1}{2}\sum [i] \cdot z^2$, where [i] is the molar concentration of the ion, and z^2 is the charge of the ion squared.

Appendix E

Analyses of the dimer-dimer interface

JavaProtein Dossier (JPD)

The dimer-dimer interface of *M. tuberculosis* DHDPS (PDB 1XXX) was analyzed with JPD (from http://sms.cbi.cnptia.embrapa.br/SMS/index_s.html), which determines the interatomic contacts established between residues belonging to two different chains at their interface.[1] There is asymmetry in the inter-subunit contacts determined with this method. For example, the hydrogen bond between R238 and L203 **(Table E.2)** occurred from chain *a* and *d*, but not *vice versa*, presumably due to the shortest atomic distance between the residues falling outside the ≤3.2 Å range **(**see **Table E.8)**.

Table E.1: Inter-subunit hydrophobic contacts (range: 2-3.8 Å) between chains *a* and *d* determined with JPD.[1]

Chain1	Residue1	Atom1	Distance (Å)	Atom2	Residue2	Chain2
A	LEU 177	CD1	3.70	CD2	LEU 198	D
A	HIS 178	CE1	3.52	CB	CYS 248	D
A	ALA 181	C	3.76	CD1	ILE 242	D
A	GLN 182	CD	3.78	CB	ALA 245	D
A	ALA 204	CB	3.40	CB	ALA 204	D
A	MET 205	CE	3.72	CG	PRO 201	D
A	MET 205	CE	3.75	CD	PRO 201	D
A	MET 205	C	3.40	CD	ARG 238	D
A	GLY 206	CA	3.71	CD	ARG 238	D
A	ILE 234	CG1	3.53	CD1	ILE 234	D
A	ILE 234	CD1	3.71	CD1	ILE 234	D
A	ARG 238	CD	3.66	C	ALA 204	D
A	ARG 238	CD	3.75	CA	MET 205	D
A	ARG 238	CD	3.39	C	MET 205	D
A	ARG 238	CZ	3.76	CA	GLY 206	D
A	ALA 245	CB	3.60	CE1	HIS 178	D
A	CYS 248	CB	3.79	CE1	HIS 178	D

Table E.2: Inter-subunit hydrogen bonds (range: 2-3.2 Å) between chains *a* and *d*, involving main chain and side chain, determined with JPD.[1]

Chain1	Residue1	Atom1	Distance (Å)	Atom2	Residue2	Chain2
A	ARG 238	NH1	3.13	O	LEU 203	D
A	ARG 238	NH1	2.94	O	ALA 204	D

Table E.3: Inter-subunit hydrogen bonds (range: 2-3.2 Å) between chains *a* and *d*, involving a single water molecule bridging main chain to side chain (main chain...water...side chain), determined with JPD.[1]

Chain1	Residue1	Atom1	Distance (Å)	H_2O	Distance (Å)	Atom2	Residue2	Chain2
A	HIS 178	ND1	2.62	17195	2.97	O	ALA 197	D
A	ALA 197	O	3.13	17895	2.62	ND1	HIS 178	D

Table E.4: Inter-subunit hydrogen bonds (range: 2-3.2 Å) between chains *a* and *d*, involving a pair of water molecules bridging main chain to side chain (main chain...water...water...side chain), determined with JPD.[1]

Chain1	Residue1	Atom1	Distance (Å)	H_2O	Distance (Å)	H_2O	Distance (Å)	Atom2	Residue2	Chain2
A	HIS 178	ND1	2.62	17195	3.16	17962	2.75	O	ALA 197	D
A	HIS 178	ND1	2.62	17195	2.84	17894	2.71	O	ASN 241	D
A	ALA 197	O	2.75	17177	3.17	17895	2.62	ND1	HIS 178	D
A	MET 205	O	2.69	17234	2.73	17862	2.76	ND2	ASN 241	D
A	ASN 241	O	2.82	17257	2.91	17895	2.62	ND1	HIS 178	D
A	ASN 241	ND2	3.03	17248	2.36	17890	2.97	O	MET 205	D

Table E.5: Inter-subunit hydrogen bonds (range: 2-3.2 Å) between chains *a* and *d*, involving a pair of side chains (side chain...side chain), determined with JPD.[1]

Chain1	Residue1	Atom1	Distance (Å)	Atom2	Residue2	Chain2
A	HIS 178	NE2	2.83	OD1	ASN 249	D
A	GLN 182	OE1	2.64	ND2	ASN 249	D
A	ASN 249	OD1	2.91	NE2	HIS 178	D
A	ASN 249	ND2	3.00	OE1	GLN 182	D

Table E.6: Inter-subunit hydrogen bonds (range: 2-3.2 Å) between chains *a* and *d*, involving a pair of water molecules bridging side chains (side chain...water...water...side chain), determined with JPD.[1]

Chain1	Residue1	Atom1	Distance (Å)	H_2O	Distance (Å)	H_2O	Distance (Å)	Atom2	Residue2	Chain2
A	HIS 178	ND1	2.62	17195	2.84	17894	3.17	OD1	ASN 241	D
A	HIS 178	ND1	2.62	17195	3.16	17962	2.76	OD1	ASN 241	D
A	ASN 241	OD1	2.67	17177	3.17	17895	2.62	ND1	HIS 178	D

Table E.7: Inter-subunit π-cation interactions (range: 2-6.0 Å) between chains a and d, determined with JPD.[1]

Chain1	Residue1	Distance (Å)	Residue2	Chain2
A	PHE 229	5.50	ARG 238	D
A	ARG 238	5.52	PHE 229	D

Protein Interfaces, Surfaces, and Assemblies (PISA) database

The dimer-dimer interface of *M. tuberculosis* DHDPS (PDB 1XXX) was also analyzed using the PISA service (from http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html) developed at the European Bioinformatics Institute.[2] This defines the interface as the protein surface area which becomes inaccessible to solvents when two chains come into contact.[2] PISA also calculates hydrogen bonds, salt bridges and disulfide bonds; however, neither of the latter two types of inter-subunit interactions were found in the dimer-dimer interface of DHDPS from *M. tuberculosis*.

Table E.8: Inter-subunit non-water-mediated hydrogen bonds (range: 2-4.0 Å) between chains *a* and *d*, determined with PISA.[2]

no.	Chain 1	Residue [atom] 1	Distance (Å)	Residue [atom] 2	Chain 2
1	D	HIS 178 [NE2]	2.91	ASN 249 [OD1]	A
2	D	GLN 182 [NE2]	3.89	ASN 249 [OD1]	A
3	D	ARG 238 [NH1]	3.43	LEU 203 [O]	A
4	D	ARG 238 [NH1]	3.27	ALA 204 [O]	A
5	D	ASN 249 [ND2]	2.64	GLN 182 [OE1]	A
6	D	GLN 182 [OE1]	3.00	ASN 249 [ND2]	A
7	D	LEU 203 [O]	3.13	ARG 238 [NH1]	A
8	D	ALA 204 [O]	2.94	ARG 238 [NH1]	A
9	D	ASN 249 [OD1]	2.82	HIS 178 [NE2]	A
10	D	ASN 249 [OD1]	3.45	GLN 182 [NE2]	A

Table E.9: Inter-subunit non-water-mediated hydrogen bonds (range: 2-4.0 Å) between chains *b* and *c*, determined with PISA.[2]

no.	Chain 1	Residue [atom] 1	Distance (Å)	Residue [atom] 2	Chain 2
1	B	HIS 178 [NE2]	2.73	ASN 249 [OD1]	C
2	B	GLN 182 [NE2]	3.35	ASN 249 [OD1]	C
3	B	ASN 249 [ND2]	3.08	GLN 182 [OE1]	C
4	B	ARG 238 [NH1]	3.11	LEU 203 [O]	C
5	B	ARG 238 [NH1]	3.01	ALA 204 [O]	C
6	B	GLN 182 [OE1]	2.61	ASN 249 [ND2]	C
7	B	LEU 203 [O]	3.11	ARG 238 [NH1]	C
8	B	ALA 204 [O]	3.16	ARG 238 [NH1]	C
9	B	ASN 249 [OD1]	3.85	GLN 182 [NE2]	C
10	B	ASN 249 [OD1]	2.80	HIS 178 [NE2]	C

Table E.10: Inter-subunit non-water-mediated hydrogen bonds (range: 2-4.0 Å) between chains *g* and *f*, determined with PISA.[2]

no.	Chain 1	Residue [atom] 1	Distance (Å)	Residue [atom] 2	Chain 2
1	G	ASN 249 [ND2]	2.77	GLN 182 [OE1]	F
2	G	ARG 238 [NH1]	3.04	LEU 203 [O]	F
3	G	ARG 238 [NH1]	2.82	ALA 204 [O]	F
4	G	HIS 178 [NE2]	2.71	ASN 249 [OD1]	F
5	G	GLN 182 [NE2]	3.65	ASN 249 [OD1]	F
6	G	GLN 182 [OE1]	2.80	ASN 249 [ND2]	F
7	G	LEU 203 [O]	3.14	ARG 238 [NH1]	F
8	G	ALA 204 [O]	3.04	ARG 238 [NH1]	F
9	G	ASN 249 [OD1]	3.46	GLN 182 [NE2]	F
10	G	ASN 249 [OD1]	2.71	HIS 178 [NE2]	F

The same hydrogen bonds were identified with PISA in all sets of dimer-dimer interactions **(Table E.8, E.9, E.10)** apparent in the asymmetric unit of the crystal structure (PDB 1XXX), including between chain *e* and *h* (not shown). There are some variation in the hydrogen bonding distances for some residues. For example, the shortest atomic distance between R238 and L203 was found to be 3.13 Å from chain *a* to *d*, and 3.43 Å *vice versa* **(Table E.8)**, whereas it was 3.11 Å from chain *b* to *c* and *vice versa* **(Table E.9)**. However, PISA uses broader criteria to define a hydrogen bond (≥4.0 Å), in contrast to JPD, which results in the identification of the same hydrogen bonds in all sets of dimer-dimer interactions in the crystal structure.

Interface residues	PYMOL		L177	H178	A181	Q182		A197			P201	L203	A204	M205				I234	I242	R238	N241				N249
	JPD		L177	H178	A181	Q182		A197	L198		P201	L203	A204	M205	G206	F229		I234	I242	R238	N241	A245		C248	N249
	PISA	A175	L177	H178	A181	Q182	A185	A197	L198	L200	P201	L203	A204	M205	G206	F229	G232	I234	I242	R238	N241	A245	P246	C248	N249
Contacts	*PYMOL*		*L177* *P201*	*D196* *A197* *N199* *L200* *N241* *N249*	*M205* *I242*	*I242* *N249*		*H178*			*L177*	*R238*	*A204* *R238*	*A181* *L200* *R238* *N241*				*I234*	*A181* *Q182*	*L203* *A204* *M205* *N241*	*H178* *L200* *M205* *R238*				*H178* *Q182*
	JPD		*L198*	*A197* *A245* *N241* *C248* *N249*	*I242*	*A245* *N249*		*H178* *N241*	*L177*		*M205*	*R238*	*A204* *R238*	*P201* *R238* *N241*	*R238*	*R238*		*I234*	*A181*	*L203* *A204* *M205* *G206* *F229*	*H178* *A197* *M205*	*H178* *Q182*		*H178*	*H178* *Q182*
	PISA			*N249*		*N249*						*R238*	*R238*							*L203* *A204*					*H178* *Q182*

Figure E.1: Comparison of interface residues and inter-subunit contacts determined by visual inspection with PYMOL,[3] by JDP,[1] and by PISA.[2] The interface residues determined using PYMOL and JDP were defined by inter-subunit contacts, whereas PISA defines the interface as the residues buried by the association of subunits. PISA determines the hydrogen bonds, but not hydrophobic contacts across the dimer-dimer interface.

References

1. Neshich, G., Rocchia, W., Mancini, A. L., Yamagishi, M. E., Kuser, P. R., Fileto, R., Baudet, C., Pinto, I. P., Montagner, A. J., Palandrani, J. F., Krauchenco, J. N., Torres, R. C., Souza, S., Togawa, R. C. & Higa, R. H. JavaProtein Dossier: a novel web-based data visualization tool for comprehensive analysis of protein structure. *Nucleic acids research* **32**, 595-601 (2004).

2. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *Journal of molecular biology* **372**, 774-97 (2007).

3. DeLano, W. L. The PyMOL molecular graphics system. (DeLano Scientific, San Carlos, CA, USA, 2006).

Appendix F

The dimer-dimer interface of *E. coli* DHDPS

The dimer-dimer interface of *M. tuberculosis* DHDPS was considerable larger than that of *E. coli* DHDPS, as quantified by the number of interface residues determined using PISA **(**24 (±1) in **Table F.2** vs. 15 (±2) in **Table F.1)**.[1] In addition, more residues are buried ≥45 % by the association of subunits in *M. tuberculosis* DHDPS, as compared to *E. coli* DHDPS.

Table F.1: PISA analysis of the dimer-dimer interface of wild-type *E. coli* DHDPS.[1] Interface residues buried ≥45 % by the association of subunits are shaded grey.

(A)

no.	Chain 1	Residue 1	H-bond	BSA[a]
1	A	GLY 165		27 %
2	A	ASN 166		2 %
3	A	LEU 167		90 %
4	A	THR 168		70 %
5	A	ASN 171	H-bond	83 %
6	A	LYS 174		4 %
7	A	GLU 175	H-bond	34 %
8	A	SER 190		68 %
9	A	ASP 193	H-bond	78 %
10	A	GLN 196	H-bond	65 %
11	A	LEU 197		100 %
12	A	GLY 198		7 %
13	A	ARG 230	H-bond	52 %
14	A	GLN 234	H-bond	37 %
15	A	MET 237		4 %

(B)

no.	Chain 1	Residue 1	H-bond	BSA[a]
1	B	GLY 165		23%
2	B	ASN 166		1%
3	B	LEU 167		89%
4	B	THR 168		68%
5	B	ASN 171	H-bond	83%
6	B	LYS 174		39%
7	B	GLU 175	H-bond	34%
8	B	ALA 189		1%
9	B	SER 190		77%
10	B	ASP 193	H-bond	80%
11	B	GLN 196	H-bond	70%
12	B	LEU 197		99%
13	B	GLY 198		5%
14	B	PHE 226		1%
15	B	ARG 230		58%
16	B	GLN 234	H-bond	37%
17	B	MET 237		3%

[a] The percentage of the total solvent-accessible surface area buried by the association of subunits.

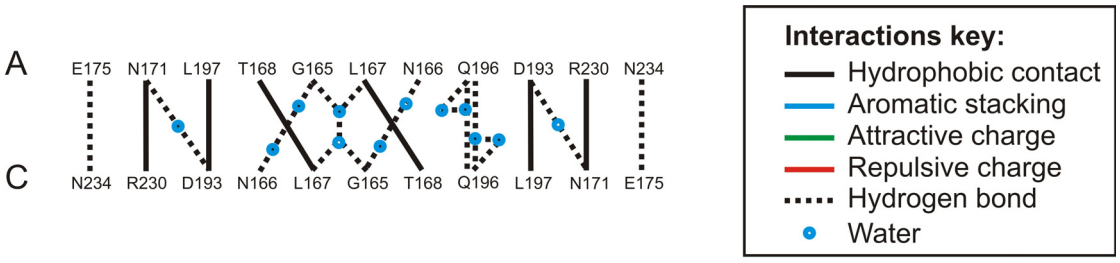


Figure F.1: JDP analysis of the dimer-dimer interface of wild-type *E. coli* DHDPS (Dr Sean Devenish *pers. comm.*).[2] This was figure produced by Dr Sean Devenish.

Table F.2: PISA analysis of the dimer-dimer interface of wild-type *M. tubeculosis* DHDPS.[1] Interface residues buried ≥45 % by the association of subunits are shaded grey.

(A)

no.	Chain 1	Residue 1	H-bond	BSA[a]
1	A	ALA 175		2%
2	A	LEU 177		88%
3	A	HIS 178	H-bond	80%
4	A	ALA 181		83%
5	A	GLN 182	H-bond	71%
6	A	ALA 185		36%
7	A	ALA 197		64%
8	A	LEU 198		54%
9	A	LEU 200		100%
10	A	PRO 201		100%
11	A	LEU 203	H-bond	25%
12	A	ALA 204	H-bond	100%
13	A	MET 205		95%
14	A	GLY 206		41%
15	A	PHE 229		61%
16	A	GLY 232		49%
17	A	ILE 234		80%
18	A	ARG 238	H-bond	54%
19	A	ASN 241		76%
20	A	ILE 242		57%
21	A	ALA 245		90%
22	A	CYS 248		54%
23	A	ASN 249	H-bond	38%

(B)

no.	Chain 1	Residue 1	H-bond	BSA[a]
1	B	ALA 175		4%
2	B	LEU 177		95%
3	B	HIS 178	H-bond	82%
4	B	ALA 181		87%
5	B	GLN 182	H-bond	62%
6	B	ALA 185		31%
7	B	ALA 197		65%
8	B	LEU 198		59%
9	B	LEU 200		100%
10	B	PRO 201		100%
11	B	LEU 203	H-bond	25%
12	B	ALA 204	H-bond	100%
13	B	MET 205		94%
14	B	GLY 206		40%
15	B	PHE 229		55%
16	B	GLY 232		36%
17	B	ILE 234		81%
18	B	ARG 238	H-bond	55%
19	B	ASN 241		77%
20	B	ILE 242		59%
21	B	ALA 245		88%
22	B	PRO 246		3%
23	B	CYS 248		56%
24	B	ASN 249	H-bond	41%

[a] The percentage of the total solvent-accessible surface area buried by the association of subunits.

The larger surface area of the dimer-dimer interface in *M. tuberculosis* DHDPS, as compared with *E. coli* DHDPS, is apparent in **Figure F.2**. The leucine residue (*E. coli*: L197, *M. tuberculosis*: M205) that successfully resulted in dimeric variants of *E. coli* DHDPS is fairly central in the interface **(Figure F.2A)**[3] and lies among those residues that are buried to a relatively high degree (≥85 %) by the association into the homotetramer quaternary structure. Similarly, the alanine residue mutated in this study (*M. tuberculosis*: A204, *E. coli*: Q196) is relatively central within the dimer-dimer interface **(Figure F.2B)** and is buried 100 % by the association of subunits.

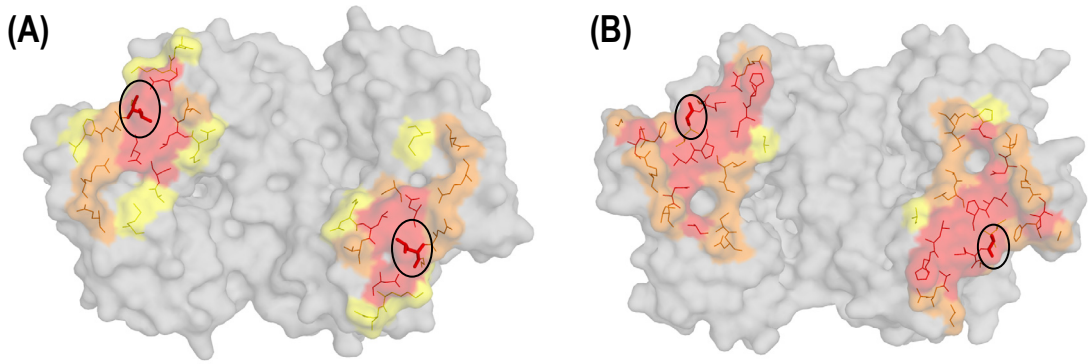


Figure F.2: **Looking down onto the dimer-dimer interface (chains *a* and *b*) of (A) *E. coli* and (B) *M. tuberculosis* DHDPS, showing the position of residues mutated to create dimeric variants (circled in black), L197 (in *E. coli* DHDPS) and A204 (in *M. tuberculosis* DHDPS), respectively. The residues contributing to the interface are shown in various shades depending on their buried surface area (5 - 44 % yellow, 45 - 84 % orange, ≥85 % red) as defined by PISA.[1]**

References

1. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *Journal of molecular biology* **372**, 774-97 (2007).

2. Neshich, G., Rocchia, W., Mancini, A. L., Yamagishi, M. E., Kuser, P. R., Fileto, R., Baudet, C., Pinto, I. P., Montagner, A. J., Palandrani, J. F., Krauchenco, J. N., Torres, R. C., Souza, S., Togawa, R. C. & Higa, R. H. JavaProtein Dossier: a novel web-based data visualization tool for comprehensive analysis of protein structure. *Nucleic acids research* **32**, 595-601 (2004).

3. Griffin, M. D., Dobson, R. C., Pearce, F. G., Antonio, L., Whitten, A. E., Liew, C. K., Mackay, J. P., Trewhella, J., Jameson, G. B., Perugini, M. A. & Gerrard, J. A. Evolution of quaternary structure in a homotetrameric enzyme. *Journal of molecular biology* **380**, 691-703 (2008).

Appendix G

Substrate inhibition at high (*S*)-ASA concentrations

Background

High concentrations of one or more of an enzyme's substrates are often found to inhibit that enzyme's activity.[1] Substrate inhibition happens when a molecule of substrate binds to the wrong enzyme form, and thus produces an inactive or dead-end complex.[2] However, this phenomenon is not commonly observed in kinetic analysis because substrate concentrations are kept at or below their physiological values.[2] The ping-pong kinetic model assigned to DHDPS is a compulsory order mechanism requiring (*S*)-ASA to bind after pyruvate. Therefore, substrate inhibition is expected at high (*S*)-ASA concentrations, due to (*S*)-ASA binding to DHDPS prior to pyruvate and forming a dead-end (*S*)-ASA-enzyme complex that has to dissociate before the reaction can proceed.

It is difficult to compare literature reports of substrate inhibition for DHDPS, or its absence, because of the different assaying and synthetic methods used by workers **(Table G.1)**. However, the majority of kinetic studies with bacterial DHDPS orthologues have not observed substrate inhibition **(Table G.1)**, probably because low concentrations of (*S*)-ASA were used for kinetic analysis to avoid such complications. The presence of substrate inhibition may have also been masked by inaccuracy, due to the lag time in some of the assaying methods (chapter two, section 2.2).

Substrate inhibition by (*S*)-ASA was observed for *E. coli* DHDPS at low concentration, giving a dissociation constant for the (*S*)-ASA-enzyme complex ($K_i^{(S)\text{-ASA}}$) of 0.3 (±0.08) mM.[3] However, it was later demonstrated that this inhibition was due to impurities in the (*S*)-ASA synthesized using Black & Wright's ozonolysis method.[4] Parallel experiments were performed contrasting (*S*)-ASA synthesized using different methods and showed that inhibition was only observed when (*S*)-ASA was prepared using ozonolysis,[5] in comparison to higher (>95 %) purity (*S*)-ASA, synthesized using the protocol by Roberts *et al.*,[6] which showed no inhibition at or below the 2 mM concentration tested.[4]

Table G.1: (*S*)-ASA inhibition observed or not during kinetic analysis of DHDPS.

Organism	Inhibition @ [(*S*)-ASA]	(*S*)-ASA synthesis	Assay method	Ref.
Bacillus licheniformis	Yes 5.22 mM	Black & Wright [5]	*o*-aminobenzaldehyde	[7]
Bacillus megaterium	None 0.01 mM	Black & Wright [5]	imidazole	[8]
Bacillus sphaericus	None 3.75 mM	Black & Wright [5]	*o*-aminobenzaldehyde	[9]
Bacillus subtilis	None 0.66 mM	Black & Wright [5]	*o*-aminobenzaldehyde	[10]
Brevibacterium lactofermentum	None 0.2 mM	Black & Wright [5]	*o*-aminobenzaldehyde	[11]
C. glutamicum	Yes 5.6 mM	Black & Wright [5]	coupled	[12]
E. coli	Yes 0.3 mM	Black & Wright [5]	coupled	[3]
E. coli	Yes 1.1 mM	Black & Wright [5]	coupled	[4]
E. coli	None 2 mM	Roberts, *et al.*[6]	coupled	[4]
Methanobacterium thermoautrophicum	None 10 mM	Black & Wright [5]	*o*-aminobenzaldehyde	[13]
Neisseria meningitidis	Yes 1.7 mM	Roberts, *et al.*[6]	coupled	[14]
Sinorhizobium meliloti	None 0.45 mM	Roberts, *et al.*[6]	imidazole	[15]
S. aureus (MRSA)	None 3.0 mM	Black & Wright [5]	coupled	[16]
S. aureus (MRSA)	Yes 2.7 mM	Roberts, *et al.*[6]	coupled	[17]
Streptomyces clavuligerus	None 0.03 mM	Black & Wright [5]	imidazole	[18]
Thermoanaerobacter tengcongensis	Yes 4.7 mM	Roberts, *et al.*[6]	coupled	[19]
T. maritima	None 1.25 mM	Roberts, *et al.*[6]	coupled	[20]

More recently, substrate inhibition has been observed at relatively low concentration of (*S*)-ASA in DHDPS enzymes from *Sinorhizobium meliloti*, *MRSA* and *Neisseria meningitidis* **(Table G.1)**.[14,15,17] Since in these studies (*S*)-ASA was synthesized using the protocol by Roberts *et al.*,[6] the inhibition observed is unlikely to be due to contaminants. Interestingly, *MRSA* DHDPS has been characterized both with and without substrate inhibition in two separate publications **(Table G.1)**.[16,17] Similar K_M values for pyruvate of 0.11 (±0.01) and 0.12 (±0.01) mM were determined, whereas a lower K_M for (*S*)-ASA was obtained from data fitted with substrate inhibition of 0.22 (±0.02) mM, compared to 0.33 (±0.03) mM from data fitted without substrate inhibition.[16,17]

Substrate inhibition due to (*S*)-ASA has recently been noted for DHDPS from *Thermoanaerobacter tengcongensis* and *C. glutamicum* occurring at comparably high (*S*)-ASA concentrations of ~5 mM **(Table G.1)**.[12,19] However, both experimental setups have alternative causes for the inhibition, as commented on by investigators.[12,19] The initial-rate data of *Thermoanaerobacter tengcongensis* DHDPS were fitted to the ping-pong (*S*)-ASA substrate inhibition model to give a $K_i^{(S)\text{-ASA}}$ of 4.7 (±0.8) mM.[19] In contrast, for *C. glutamicum* DHDPS, even though substrate inhibition was observed at ≥5.6 mM (*S*)-ASA, the initial-rate data were fitted to the ping-pong model without substrate inhibition.[12] This was justified by

the normal Michaelis-Menten type saturation curves observed within the (*S*)-ASA concentration ranges of 0.14 mM and 2.8 mM used for kinetic characterization.[12]

M. tuberculosis DHDPS

Kinetic analysis of wild-type *M. tuberculosis* DHDPS investigated the (*S*)-ASA concentration range from 0.03 to 3.0 mM and showed no substrate inhibition (chapter two, section 2.4.4). However, the kinetic analysis of DHDPS-A204R required higher (*S*)-ASA concentrations and initial investigation showed curvature suggestive of substrate inhibition. De-activation of the enzyme due to a decrease in pH due to the acidity of the (*S*)-ASA solution was eliminated by using a high concentration of HEPES buffer in the coupled assay protocol, which neutralized the acidity even at the highest concentrations of (*S*)-ASA investigated. Thus pseudo-single substrate kinetic experiments were performed for DHDPS-A204R and wild-type DHDPS **(Figure G.1)**.

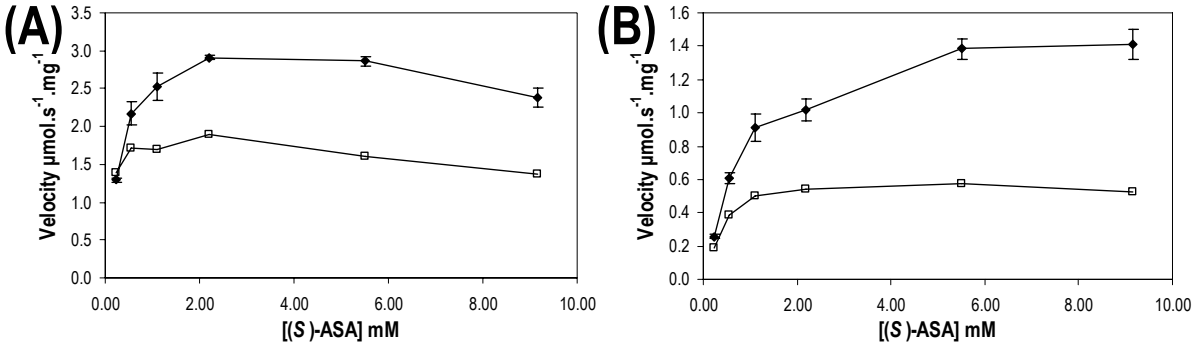


Figure G.1: The effect of (*S*)-ASA concentration on DHDPS activity, as measured by the coupled assay for (A) wild-type and (B) A204R and *M. tuberculosis* DHDPS. The pyruvate in the assay mixture was held constant at high (◆ 4.0 mM) or low concentrations (□ 0.15 mM for A204R and 0.05 mM for wild-type). Duplicate measurements were recorded for the data points with error bars to show the standard deviation.

Substrate inhibition at high (*S*)-ASA concentrations is more obvious in wild-type *M. tuberculosis* DHDPS with the activity starting to decrease at concentrations above 3 mM **(Figure G.1A)**; although, this is above the concentration range used in the characterization of the wild-type enzyme. A slight decrease in activity with DHDPS-A204R is observed only at low pyruvate concentration **(Figure G.1B)**. The protection against inhibition at higher pyruvate concentrations is consistent with the competitive inhibition expected for the ping-

pong model. A normal Michaelis-Menten type saturation curve for DHDPS-A204R is observed within the (*S*)-ASA concentration range of 0.22 to 5.5 mM used for kinetic characterization, indicating that the kinetic data can be fitted to the ping-pong model without substrate inhibition.

The K_M value of 1.1 (±0.1) mM for (*S*)-ASA determined from fitting the ping-pong model without substrate inhibition to the initial-rate data of DHDPS-A204R data is approximately triple the K_M value previously determined for wild-type of 0.43 (±0.02) mM. This indicates that DHDPS-A204R has a lower affinity for (*S*)-ASA, consistent with the difference in observed substrate inhibition at high (*S*)-ASA concentrations between the variant and wild-type enzyme.

References

1. Eisenthal, R. & Danson, M. J. *Enzyme assays: a practical approach* (IRL Press at Oxford University Press, Oxford [England]; New York, 1992).

2. Cornish-Bowden, A. *Fundamentals of enzyme kinetics* (Butterworths, London ; Boston, 1979).

3. Karsten, W. E. Dihydrodipicolinate synthase from *Escherichia coli*: pH dependent changes in the kinetic mechanism and kinetic mechanism of allosteric inhibition by L-lysine. *Biochemistry* **36**, 1730-9 (1997).

4. Dobson, R. C., Gerrard, J. A. & Pearce, F. G. Dihydrodipicolinate synthase is not inhibited by its substrate, (*S*)-aspartate β-semialdehyde. *The Biochemical journal* **377**, 757-62 (2004).

5. Black, S. & Wright, N. G. Aspartic β-semialdehyde dehydrogenase and aspartic β-semialdehyde. *The Journal of biological chemistry* **213**, 39-50 (1955).

6. Roberts, S. J., Morris, J. C., Dobson, R. C. & Gerrard, J. A. The preparation of *(S)*-aspartate semi-aldehyde appropriate for use in biochemical studies. *Bioorganic & medicinal chemistry letters* **13**, 265-7 (2003).

7. Stahly, D. P. Dihydrodipicolinic acid synthase of *Bacillus licheniformis*. *Biochimica et biophysica acta* **191**, 439-51 (1969).

8. Webster, F. H. & Lechowich, R. V. Partial purification and characterization of dihydrodipicolinic acid synthetase from sporulating *Bacillus megaterium*. *Journal of bacteriology* **101**, 118-26 (1970).

9. Bartlett, A. T. M. & White, P. J. Regulation of the enzymes of lysine biosynthesis in *Bacillus sphaericus* NCTC 9602 during vegetative growth. *Journal of general microbiology* **132**, 3169-77 (1986).

10. Yamakura, F., Ikeda, Y., Kimura, K. & Sasakawa, T. Partial purification and some properties of pyruvate-aspartic semialdehyde condensing enzyme from sporulating *Bacillus subtilis*. *Journal of biochemistry* **76**, 611-21 (1974).

11. Tosaka, O., Ishihara, M., Morinaga, Y. & Takinami, K. Mode of conversion of asparto β-semialdehyde to L-threonine and L-lysine in *Brevibacterium lactofermentum*. *Agricultural and biological chemistry* **43**, 265-70 (1979).

12. Rice, E. A., Bannon, G. A., Glenn, K. C., Jeong, S. S., Sturman, E. J. & Rydel, T. J. Characterization and crystal structure of lysine insensitive *Corynebacterium glutamicum* dihydrodipicolinate synthase (cDHDPS) protein. *Archives of biochemistry and biophysics* **480**, 111-21 (2008).

13. Bakhiet, N., Forney, F., Stahly, D. & Daniels, L. Lysine biosynthesis in *Methanobacterium thermoautotrophicum* is by the diaminopimelic acid pathway. *Current microbiology* **10**, 195-8 (1984).

14. Devenish, S. R., Huisman, F. H., Parker, E. J., Hadfield, A. T. & Gerrard, J. A. Cloning and characterisation of dihydrodipicolinate synthase from the pathogen *Neisseria meningitidis*. *Biochimica et biophysica acta* **1794**, 1168-74 (2009).

15. Tam, P. H., Phenix, C. P. & Palmer, D. R. MosA, a protein implicated in rhizopine biosynthesis in *Sinorhizobium meliloti* L5-30, is a dihydrodipicolinate synthase. *Journal of molecular biology* **335**, 393-7 (2004).

16. Girish, T. S., Sharma, E. & Gopal, B. Structural and functional characterization of *Staphylococcus aureus* dihydrodipicolinate synthase. *FEBS letters* **582**, 2923-30 (2008).

17. Burgess, B. R., Dobson, R. C., Bailey, M. F., Atkinson, S. C., Griffin, M. D., Jameson, G. B., Parker, M. W., Gerrard, J. A. & Perugini, M. A. Structure and evolution of a novel dimeric enzyme from a clinically-important bacterial pathogen. *The Journal of biological chemistry* **283**, 27598-603 (2008).

18. Mendelovitz, S. & Aharonowitz, Y. Regulation of cephamycin C synthesis, aspartokinase, dihydrodipicolinic acid synthetase, and homoserine dehydrogenase by aspartic acid family amino acids in *Streptomyces clavuligerus*. *Antimicrobial agents and chemotherapy* **21**, 74-84 (1982).

19. Wolterink-van Loo, S., Levisson, M., Cabrieres, M. C., Franssen, M. C. & van der Oost, J. Characterization of a thermostable dihydrodipicolinate synthase from *Thermoanaerobacter tengcongensis*. *Extremophiles* **12**, 461-9 (2008).

20. Pearce, F. G., Perugini, M. A., McKerchar, H. J. & Gerrard, J. A. Dihydrodipicolinate synthase from *Thermotoga maritima*. *The Biochemical journal* **400**, 359-66 (2006).

Appendix H

Analysis of pyruvate effect *via* biophysical methods

The effect of pyruvate on the oligomeric state of DHDPS-A204R was examined by two biophysical techniques. Experiments performed with wild-type *M. tuberculosis* DHDPS allow monitoring of the inaccuracies that may be involved in these experimental setups.

Wild-type *M. tuberculosis* DHDPS, like DHDPS-A204R, was purified using a preparative gel-filtration Superdex 200 16/60 column (GE Healthcare) pre-equilibrated with purification buffer containing 10 mM pyruvate. From the four-point calibration plot that was generated **(Figure H.1B),** the symmetrical elution peak of the wild-type enzyme **(Figure H.1A)** was determined to correspond to a molecular mass of 117 kDa.

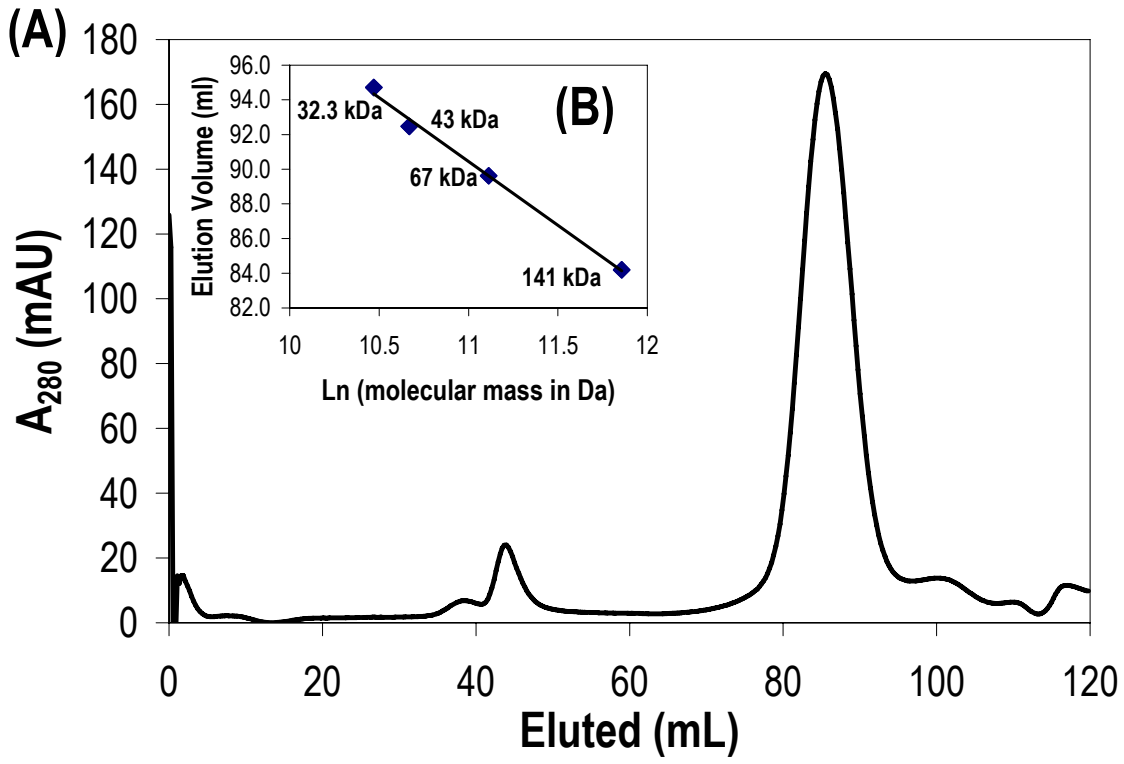


Figure H.1: **(A) Gel filtration chromatography of *M. tuberculosis* DHDPS, performed at 4 ºC in purification buffer [20 mM Tris-HCl, 250 mM NaCl, 5 % glycerol, 2 mM β-ME, 10 mM pyruvate] gave an elution peak with a maximum at 85.49 mL. (B) BSA, ovalbumin, and alcohol dehydrogenase calibration standards (◆) were fitted to a linear equation.**

This value of 117 kDa for wild-type *M. tuberculosis* DHDPS is consistent with its tetrameric quaternary structure and the solution molar masses of 116 and 105 kDa determined with analytical ultracentrifugation and analytical gel-filtration, respectively (in chapter two, section 2.5). Since, the preparative gel-filtration setup gave a similar molar mass as other methods, differences observed when analyzing DHDPS-A204R could not be attributed to the effect of the buffer components or temperature on the preparative gel-filtration column itself.

Sedimentation velocity experiments with wild-type *M. tuberculosis* DHDPS were performed at 40000 rpm and at 20 ºC, in the storage buffer [20 mM Tris.HCl, 250 mM NaCl, 5 % glycerol, 2 mM β-mercaptoethanol (β-ME), 10 mM pyruvate, pH 8.0]. Due to the limitations of the SEDNTERP program partial specific volume ($\bar{v}$), solvent density (ρ) and viscosity (η) were estimates (of 0.7402 $mL.g^{-1}$, 1.023 $g.mL^{-1}$ and 1.193 cP, respectively), not taking into account some possible effects due to buffer components.

The data were collected every 6 minutes and fitted to a continuous sedimentation coefficient [*c(s)*] distribution model, using SEDFIT,[1] which gave a main peak of 5.3 S **(Figure H.2B)**. The second peak at 3.0 S likely corresponds to the contaminant that absorbs more highly at 280 nm than 230 nm, which is discussed in detail in chapter two, section 2.5.2. Plots of raw data overlaid with the calculated fits were produced for visual inspection of the analysis **(Figure H.2A)**. The fit was less than ideal, as indicated by the high Z-test value of 20.49; however, the randomly distributed residuals and low r.m.s.d. indicate the relative goodness of the fit **(Figure H.2)**. The frictional ratio (f/f_0) of 1.28 is significantly different to previous f/f_0 of 1.23 and 1.22 extracted from fitting data of wild-type *M. tuberculosis* DHDPS in 20 mM Tri.HCl, 150 mM NaCl, pH 8.0 (in chapter two, section 2.5.2, Figures 2.12 and 2.13) and may reflect effects due to the buffer components not present during these other experiments.

The sedimentation coefficient is relatively unaffected by the frictional ratio (f/f_0),[2] and thus can be provide insight into the oligomeric species present regardless of possible inaccuracies in the f/f_0. HYDROPRO predicts a sedimentation coefficient of 5.5 S for tetrameric DHDPS, indicating that the main peak observed of 5.3 S corresponds to a tetramer, consistent with the quaternary structure of *M. tuberculosis* DHDPS. In addition, when this sedimentation coefficient was standardized to conditions corresponding to pure water at 20 °C, using SEDNTERP,[3] a value of 6.8 S was obtained, which is similar to 6.5 S for $s°_{20,w}$ found for wild-type *M. tuberculosis* DHDPS in 20 mM Tris.HCl, 150 mM NaCl, pH 8.0 (in chapter

two, section 2.5.2). This demonstrates that estimations of solution and protein properties for SV analysis in storage buffer do not obscure the determination of oligomeric state, even though there may be some inaccuracy in the frictional ratio.

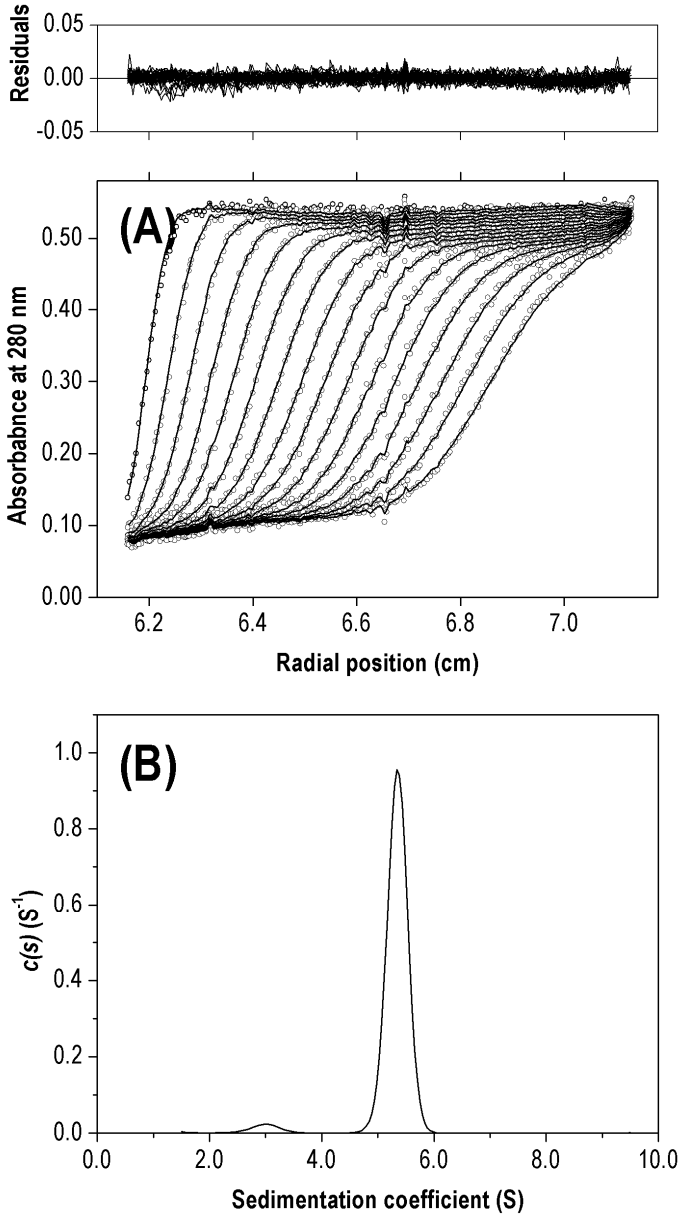


Figure H.2: **Sedimentation velocity analysis of *M. tuberculosis* DHDPS (1.0 mg.mL^{-1}) in storage buffer [20 mM Tris.HCl, pH 8.0, 250 mM NaCl, 5 % glycerol, 2 mM β-ME, and 10 mM pyruvate] at 20 ºC. (A) Absorbance at 280 nm (○) is plotted at time intervals of 12 minutes and overlaid with the nonlinear least-squares fit (solid line) to a [*c(s)*] model.[1] (B) The fit was obtained using a resolution of 200 species with $\bar{v}$ = 0.7402 mL.g^{-1}, ρ = 1.023 g.mL^{-1}, η = 1.193 cP and f/f_0 = 1.28437, and TI noise was removed. The r.m.s.d. and Z-test for the fit were 0.0042 and 20.49, respectively.**

References

1. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modelling. *Biophysical journal* **78**, 1606-19 (2000).

2. Balbo, A. & Schuck, P. Analytical ultracentrifugation in the study of protein self-association and heterogeneous protein-protein interactions in *Protein-protein interactions: a molecular cloning manual* (ed. Golemis, E.) 253-78 (Cold Spring Harbor Laboratory Press, New York, 2005).

3. Laue, T. M., Shah, D. B., Ridgeway, T. M. & Pelletier, S. L. *Analytical ultracentrifugation in biochemistry and protein science* (The Royal Society of Chemistry, Cambridge, 1992).

Appendix I

Crystallization of DHDPS-A204R

Crystallization of *M. tuberculosis* DHDPS-A204R was performed by Linda Schuldt, under the supervision of Dr Manfred Weiss, from the EMBL Hamburg Outstation, Germany. Purified DHDPS-A204R was concentrated to 10 $mg.mL^{-1}$ using an ultrafiltration spin column (Vivaspin 15, molecular cut-off >10000 kDa, range 2-8 mL) in our lab and sent to Germany.

Our collaborator, Linda Schuldt, screened for initial crystallization conditions and crystals were observed in the several conditions, all of which contained 2.0 M ammonium sulfate and various buffers at different pH values at room temperature (**Figure I.1A**). Optimization was carried out using the sitting-drop vapor diffusion method in Greiner 96-well plates. 200 nL of protein solution and 200 nL of reservoir solution were equilibrated against 30 µL of reservoir solution. Compact crystals grew out of 2.0 M ammonium sulfate and 100 mM sodium acetate pH 5.5 at room temperature within 4 weeks to a maximum size of 100 µm × 40 µm × 20 µm (**Figure I.1B**). These crystals diffracted X-rays up to a resolution of about 2.0 Å.

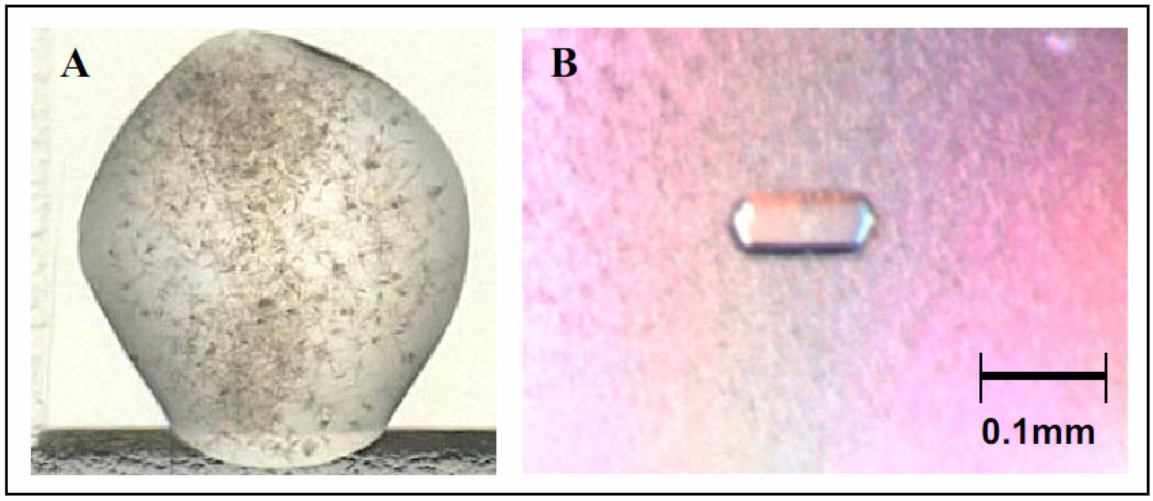


Figure I.1: (A) Crystals of *M. tuberculosis* DHDPS-A204R in the initial conditions for crystallization of 2.0 M $(NH_4)_2SO_4$ and 100 mM Tris.HCl, pH 8.5. (B) A crystal formed in optimized conditions of 2.0 M $(NH_4)_2SO_4$ and 100 mM sodium acetate pH 5.5.

Appendix J

Journal publication

Crystal structure and kinetic study of dihydrodipicolinate synthase from *Mycobacterium tuberculosis*

Kefala, G., Evans, G. L., Griffin, M. D., Devenish, S. R., Pearce, F. G., Perugini, M. A., Gerrard, J. A., Weiss, M. S. & Dobson, R. C.

BJ www.biochemj.org

BJ Structure

Biochem. J. (2008) **411**, 351–360 (Printed in Great Britain) doi:10.1042/BJ20071360

online data

Crystal structure and kinetic study of dihydrodipicolinate synthase from *Mycobacterium tuberculosis*

Georgia KEFALA*[1,2], Genevieve L. EVANS†[2], Michael D. W. GRIFFIN‡, Sean R. A. DEVENISH†, F. Grant PEARCE†, Matthew A. PERUGINI‡, Juliet A. GERRARD†, Manfred S. WEISS*[3] and Renwick C. J. DOBSON‡[3]

*EMBL Hamburg Outstation, c/o DESY, Notkestrasse 85, D-22603 Hamburg, Germany, †School of Biological Sciences, University of Canterbury, Private Bag 4800, Christchurch, New Zealand, and ‡Department of Biochemistry and Molecular Biology, Bio21 Molecular Science and Biotechnology Institute, 30 Flemington Road, University of Melbourne, Melbourne, Victoria 3010, Australia

The three-dimensional structure of the enzyme dihydrodipicolinate synthase (KEGG entry Rv2753c, EC 4.2.1.52) from *Mycobacterium tuberculosis* (*Mtb*-DHDPS) was determined and refined at 2.28 Å (1 Å = 0.1 nm) resolution. The asymmetric unit of the crystal contains two tetramers, each of which we propose to be the functional enzyme unit. This is supported by analytical ultracentrifugation studies, which show the enzyme to be tetrameric in solution. The structure of each subunit consists of an N-terminal $(\beta/\alpha)_8$-barrel followed by a C-terminal α-helical domain. The active site comprises residues from two adjacent subunits, across an interface, and is located at the C-terminal side of the $(\beta/\alpha)_8$-barrel domain. A comparison with the other known DHDPS structures shows that the overall architecture of the active site is largely conserved, albeit the proton relay motif comprising Tyr^{143}, Thr^{54} and Tyr^{117} appears to be disrupted. The kinetic parameters of the enzyme are reported: $K_M^{ASA} = 0.43 \pm 0.02$ mM, $K_M^{pyruvate} = 0.17 \pm 0.01$ mM and $V_{max} = 4.42 \pm 0.08$ $\mu mol \cdot s^{-1} \cdot mg^{-1}$. Interestingly, the V_{max} of *Mtb*-DHDPS is 6-fold higher than the corresponding value for *Escherichia coli* DHDPS, and the enzyme is insensitive to feedback inhibition by (*S*)-lysine. This can be explained by the three-dimensional structure, which shows that the (*S*)-lysine-binding site is not conserved in *Mtb*-DHDPS, when compared with DHDPS enzymes that are known to be inhibited by (*S*)-lysine. A selection of metabolites from the aspartate family of amino acids do not inhibit this enzyme. A comprehensive understanding of the structure and function of this important enzyme from the (*S*)-lysine biosynthesis pathway may provide the key for the design of new antibiotics to combat tuberculosis.

Key words: diaminopimelate pathway, dihydrodipicolinate synthase (DHDPS), lysine biosynthesis, *Mycobacterium tuberculosis*, succinylase branch.

INTRODUCTION

Mycobacterium tuberculosis (*Mtb*), the causative agent of tuberculosis, causes more deaths than any other bacterium [1,2]. The increase in tuberculosis cases worldwide, particularly among immunocompromised individuals, combined with the increase in multidrug-resistant strains, highlights the need for new antituberculosis drugs [3].

Mycobacterial cell walls are characterized by an unusually high DAP (diaminopimelic acid) content. DAP, an intermediate of the (*S*)-lysine biosynthetic pathway, is a constituent of the short peptide bridges that cross-link peptidoglycan polymer chains. Consequently, the absence of DAP results in cell lysis and death, as has been demonstrated in gene-knockout experiments with *Mycobacterium smegmatis* [4]. The process of cell wall assembly and the biosynthesis of cell wall components have long been accepted as targets for antibiotic design and many existing antibiotics inhibit key steps therein [5]. Therefore an inhibitor that inactivates any of the enzymes that are unique to the (*S*)-lysine biosynthetic pathway, preventing the synthesis of these crucial metabolites, would be a very effective antibiotic [6]. Additionally, the absence of the (*S*)-lysine pathway in mammals means that inhibitors of this pathway would not be expected to have mammalian toxicity [6].

The enzyme DHDPS (dihydrodipicolinate synthase) (KEGG entry Rv2753c; EC 4.2.1.52) catalyses the first unique reaction of (*S*)-lysine biosynthesis [7]: an aldol condensation between (*S*)-ASA [(*S*)-aspartate β-semialdehyde] and pyruvate (Scheme 1) [8]. The product of the reaction is the unstable heterocycle HTPA [(4*S*)-4-hydroxy-2,3,4,5-tetrahydro-(2*S*)-dipicolinate], which is thought to undergo a non-enzymatic dehydration to (*S*)-2,3-dihydrodipicolinate, the substrate of the next enzyme in the (*S*)-lysine biosynthetic pathway, DHDPR (dihydrodipicolinate reductase) [9,10]. The proposed Ping Pong mechanism of DHDPS involves Schiff base formation between pyruvate and an active-site lysine residue, followed by the release of water and then enamine formation [7], which has been suggested to be essentially irreversible [11]. (*S*)-ASA then binds to the stable substituted enzyme form, and this event is followed by the release of the product and the regeneration of the free enzyme [10]. A catalytic triad, involving residues from two monomers, was shown to have a crucial role in catalysis; the current hypothesis is that these three residues are involved in the transport of protons from the active site to bulk solvent [12].

All DHDPS enzymes that have had both their structure solved and function confirmed are homotetramers. The three-dimensional structures of DHDPS from *Escherichia coli* [10,13,14], *Nicotiana sylvestris* [9], *Thermotoga maritima* [15] and *Bacillus*

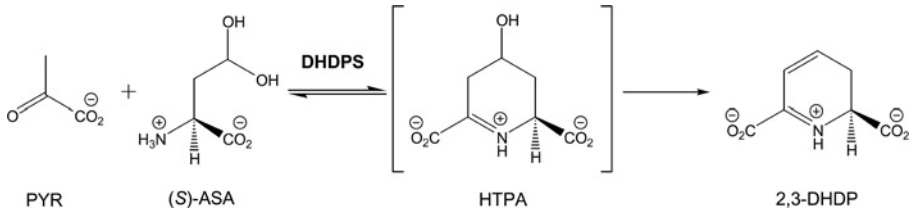
Abbreviations used: DAP, diaminopimelate; DHDPR, dihydrodipicolinate reductase; DHDPS, dihydrodipicolinate synthase; DTT, dithiothreitol; HTPA, (4*S*)-4-hydroxy-2,3,4,5-tetrahydro-(2*S*)-dipicolinate; *Mtb*, *Mycobacterium tuberculosis*; (*S*)-ASA, (*S*)-aspartate β-semialdehyde.

[1] Present address: The Salk Institute, 10010 North Torrey Pines Road, La Jolla, CA 92037, U.S.A.

[2] These authors contributed equally to this work.

[3] Correspondence may be addressed to either of these authors (email msweiss@embl-hamburg.de or rdobson@unimelb.edu.au).

The structural co-ordinates for *Mycobacterium tuberculosis* dihydrodipicolinate synthase will appear in the Protein Data Bank under accession code 1XXX.



Scheme 1 Reaction catalysed by DHDPS

anthracis [16], as well as the structures of five point mutants of the *E. coli* DHDPS [12,17], have been determined.

Characterized DHDPS enzymes display a quaternary structure best described as a dimer of tight dimers, with many interactions between the two monomers to form the tight dimer, and, in the case of the bacterial enzymes, relatively fewer interactions between the tight dimers that form the tetramer [13]. Recently, the structure of a putative DHDPS from *Agrobacterium tumefaciens* was determined to be a hexamer (PDB code 2HMC), although the function of this enzyme (and its annotation as a DHDPS) has yet to be confirmed. Despite originating from different kingdoms of life, the structure of the monomeric subunits of *E. coli* and the plant species *N. sylvestris* DHDPS enzymes are strikingly similar.

Intriguingly, the quaternary structure of the plant and bacterial species is quite different. The arrangement of the tight dimer is the same, but the arrangement of the two tight dimers to form the tetramer is different. The altered conformation can perhaps be explained by the proposed mechanism of allosteric inhibition. The activity of DHDPS is moderately inhibited by (*S*)-lysine in some bacteria, such as *E. coli* [10], whereas, in plants, a greater degree of inhibition is observed [9]. The structures of DHDPS from both *E. coli* and *N. sylvestris* complexed with (*S*)-lysine show that one (*S*)-lysine molecule is bound to each monomer, but, at the same time, is co-ordinated by residues from both monomeric units at the tight-dimer interface [9,10]. Upon (*S*)-lysine binding, a significant conformational change occurs in the *N. sylvestris* DHDPS structure [9], which is in striking contrast with the situation in the *E. coli* DHDPS, where relatively few residues shift upon (*S*)-lysine binding [13].

The focus of the present study is to understand the structure–function relationship in *Mtb*-DHDPS, which will underpin the rational design of antimicrobials. In this paper, we present an in-depth structural and kinetic study of the 'toolkit' used by *Mtb*-DHDPS in catalysis.

EXPERIMENTAL

Materials

Unless otherwise stated, all chemicals were obtained from Sigma Chemical Co., GE Biosciences or Invitrogen, and all enzyme manipulations were carried out at 6 °C or on ice. Protein concentration was measured by a modification of the Bradford method, with improved linearity over a broader range of concentrations [18] using BSA as a standard.

Overexpression and purification of *Mtb*-DHDPS

For the structural studies, the cloning, expression, purification and crystallization of *Mtb*-DHDPS has been described previously [19]. In brief, the *M. tuberculosis dap*A gene (Rv2753c) was cloned and expressed in *E. coli* cells and purified to homogeneity by affinity and size-exclusion chromatography.

For functional studies, the pETM11 plasmid containing the *M. tuberculosis dap*A gene was transformed into *E. coli* BL21(DE3) cells containing the pGroESL plasmid, which codes for the GroEL and GroES chaperones [20]. This step increased the amount of *Mtb*-DHDPS isolated. Cells were grown overnight at 37 °C and induced with IPTG (isopropyl β-D-thiogalactoside) for 3 h. Both the extraction and elution buffers contained 10 mM pyruvate, which has been used previously to stabilize *E. coli* DHDPS [11]. Following elution from a 5 ml Ni-charged column, fractions containing DHDPS activity (determined using the *o*-aminobenzaldehyde assay [21]) were incubated overnight with recombinant TEV (tobacco etch virus) protease [1 mM DTT (dithiothreitol) and 5 mM EDTA] to remove the His_6 tag, then dialysed or exchanged into a storage buffer [20 mM Tris/HCl, 2 mM 2-mercaptoethanol and 5 % (v/v) glycerol (pH 8)]. Any remaining contaminants were removed by gel filtration, and the active peak, which eluted with a molecular mass of approx. 120 kDa, consistent with a homotetramer, was collected. The purified *Mtb*-DHDPS was analysed by SDS/PAGE (4–12 % gel) and blue native PAGE (4–16 % gel) (NuPAGE Bis-Tris gel; Invitrogen) [22], and stored in storage buffer containing 250 mM NaCl at 6 or 22 °C for several weeks without detectable degradation.

Crystallization, structure solution and refinement

Crystals of *Mtb*-DHDPS were grown in the presence of 28 % (w/v) PEG [poly(ethylene glycol)] 4000, 170 mM $MgCl_2$ and 100 mM Tris/HCl (pH 8.5). Crystals were flash-cooled in a nitrogen stream at −173 °C, using 20 % (v/v) MPD (2-methyl-2,4-pentanediol) in reservoir solution as cryoprotectant. Diffraction data were collected on the XRD (X-ray diffraction) beamline at the ELETTRA synchrotron (Trieste, Italy) using a MAR CCD (charge-coupled device) (165 mm) detector. The crystals belong to the primitive monoclinic space group $P2_1$ with the following unit cell parameters: $a = 94.79$ Å (1 Å = 0.1 nm), $b = 87.37$ Å, $c = 139.85$ Å, $\beta = 107.78°$ [19]. Based on the self-rotation function, it was concluded that the asymmetric unit of the crystals contained two tetramers exhibiting D_2 symmetry. The structure of *Mtb*-DHDPS was solved by molecular replacement using a single monomer of *T. maritima* DHDPS (PDB code 1O5K [15]) as a search model. The molecular replacement solution was then subjected to rigid-body, positional and *B*-factor refinement protocols as implemented in CNS [23]. At this point the free R-factor had fallen to 46 %, indicating the correctness of the solution. The correct amino acid sequence was introduced for one monomer using GUISIDE [24] and the co-ordinates for the remaining seven monomers were generated utilizing the non-crystallographic symmetry. The model was improved further using iterative manual model corrections using the program O [25] and refinement in REFMAC5 [26]. Non-crystallographic symmetry restraints were used throughout the refinement. Water molecules were placed using the program ARP/wARP [24]. The refinement statistics are summarized in Table 1. The quality

Table 1 Refinement and model statistics for *Mtb*-DHDPS

R.m.s.d, root mean square deviation.

Parameter	Value
Space group	$P2_1$
Cell dimensions	$a = 94.79$ Å, $b = 87.37$ Å, $c = 139.85$ Å, $\beta = 107.78°$
Resolution limits (Å)	99–2.28 (2.32–2.28)
Completeness (%)	99.7 (95.0)
Resolution limits for refinement (Å)	30.0–2.28 (2.34–2.28)
Data cut-off [$F/\sigma(F)$]	0.0
Number of reflections	99295
Working set	97308
Test set	1987
R_{cryst} (%)	14.9 (16.5)
R_{free} (%)	21.5 (23.5)
Number of atoms	18793
Protein	17072
Mg^{2+} ions	8
Cl^- ions	8
Water molecules	1587
DTT molecules	8
R.m.s.d.	
Bonds (Å)	0.018
Angles (°)	1.63
Average *B* factors	
Protein (Å^2)	29.4
Ions (Å^2)	36.3
DTT (Å^2)	44.4
Water (Å^2)	30.3
Ramachandran plot	
Most favoured (%)	90.4
Additionally allowed (%)	9.2

of the model was checked using the program PROCHECK [27]. The Figures illustrating structural details were prepared using the program PyMOL (DeLano Scientific), unless stated otherwise. The refined co-ordinates, as well as the corresponding structure factor amplitudes, were deposited with the PDB under the accession number 1XXX.

Sequence alignments were derived from the three-dimensional alignments of the DHDPS structures, which were carried out using the programs ALIGN [29], STAMP [30] and LSQKAB [31]. Buried surface areas were calculated with DSSP [31], the EBI PISA server [32] and the Protein–Protein Interaction server (http://www.biochem.ucl.ac.uk/bsm/PP/server).

Enzyme activity assay

The activity of DHDPS was studied using a coupled assay as described previously [12], with the following components: pyruvate, (*S*)-ASA, DHDPR, NADPH and Hepes buffer. (*S*)-ASA was synthesized using the methods of Roberts et al. [33] and was of high quality (> 95 %) as judged by ^{1}H-NMR. Control assays were performed to ensure the absence of contaminating NADPH-utilizing enzymes and to ensure an excess of DHDPR. The coupled assay follows the consumption of NADPH by the next enzyme in the (*S*)-lysine biosynthetic pathway, DHDPR [12,34,35]. Coupling enzyme was purified from *T. maritima*. All components were incubated for 15 min before initiation of the reaction by the addition of DHDPS. The temperature was kept constant at 30 °C by the use of a circulating water bath and the pH was maintained at 8.0 by 100 mM Hepes buffer (from a stock of 200 mM Hepes, pH 8.0, at 22 °C). Selwyn's test [36] was performed to ensure that enzymes and substrates were stable over the course of the assay (results not shown). The initial rate of NADPH consumption was measured in triplicate, and was reproducible ($\pm$ 10 %). The data were fitted with the appropriate kinetic models using the program ENZFITTER (Biosoft).

Enzyme stability

The optimum pH for enzyme activity was determined using a series of 20 mM buffers (Mes, Hepes or Bicine) covering a pH range of 6–9 at 30 °C with ionic strength of 0.15 M (adjusted by the addition of NaCl). The optimum ionic strength for the enzymatic reaction was determined using a series of 20 mM Hepes buffers with various NaCl concentrations, corrected for the amount of HCl/NaOH added to bring the pH to 8.25 at 30 °C. For heat-stability assays, DHDPS from *M. tuberculosis* or *E. coli* was buffer-exchanged into 20 mM Hepes (pH 8.25, at 30 °C with an ionic strength of 0.15 M) before incubation at various temperatures (30–100 °C) using a solid heat block. Aliquots of 10 μl were taken after an incubation of 5 min, at a range of temperatures (30–100 °C), or at 70 °C with a range of incubation times (0–60 min), and stored on ice or added directly to initiate the coupled assay. The heat stability of DHDPS from *M. tuberculosis* and *E. coli* was also analysed with a thermal melt in 15 mM Tris/HCl solutions (ionic strength of 0.15 M) containing 10× Sypro Orange (Molecular Probes) with protein unfolding monitored by fluorescence emission at 575 nm [37]. The wavelength of excitation was 490 nm, and the plate was heated from 20 to 90 °C by an iQ5 Real Time PCR Detection System (Bio-Rad Laboratories).

Analytical ultracentrifugation

Sedimentation velocity experiments were performed with purified *Mtb*-DHDPS using an An-60 Ti four-hole rotor in a Beckman Coulter Model XL-I analytical ultracentrifuge equipped with a UV–visible absorbance optical system. Double-sector cells with quartz windows were loaded with 400 μl of reference (20 mM Tris/HCl and 0.15 M NaCl, pH 8.0) and 380 μl of sample (0.055 mg $\cdot$ ml^{-1}) and centrifuged at 40 000 rev./min and 20 °C. Radial absorbance data were collected at 230 nm every 6 min without averaging. Data were fitted to a continuous size-distribution model using the program SEDFIT [38]. The program SEDNTERP [39] was used to determine the partial specific volume ($\bar{v}$) of the sample (0.7402 ml $\cdot$ g^{-1}), buffer density (1.005 g $\cdot$ ml^{-1}), buffer viscosity (1.021 cP), and the standardized sedimentation coefficient ($S^0_{20,w}$). The predicted standardized sedimentation values for three different geometries of tetramer were calculated using eqn (1) [40], where M is the predicted molar mass of the tetramer, and F is the geometric factor.

$$s^0_{20,w} = \frac{0.01M^{2/3}(1-\overline{v}\rho)}{\overline{v}^{1/3}}F \quad (1)$$

RESULTS AND DISCUSSION

Overall structure

Mtb-DHDPS is a tetramer comprising four identical subunits arranged in D_2 symmetry (Figure 1A). Each monomer (300 amino acid residues) comprises an N-terminal $(\beta/\alpha)_8$-barrel domain (residues 1–233) and a C-terminal domain (residues 234–300) consisting of three α-helices (Figures 1B–1D). The residues responsible for substrate binding and catalysis are located in the $(\beta/\alpha)_8$-barrel domain. The crystallographic asymmetric unit contains two tetramers of the enzyme. Each tetramer can be

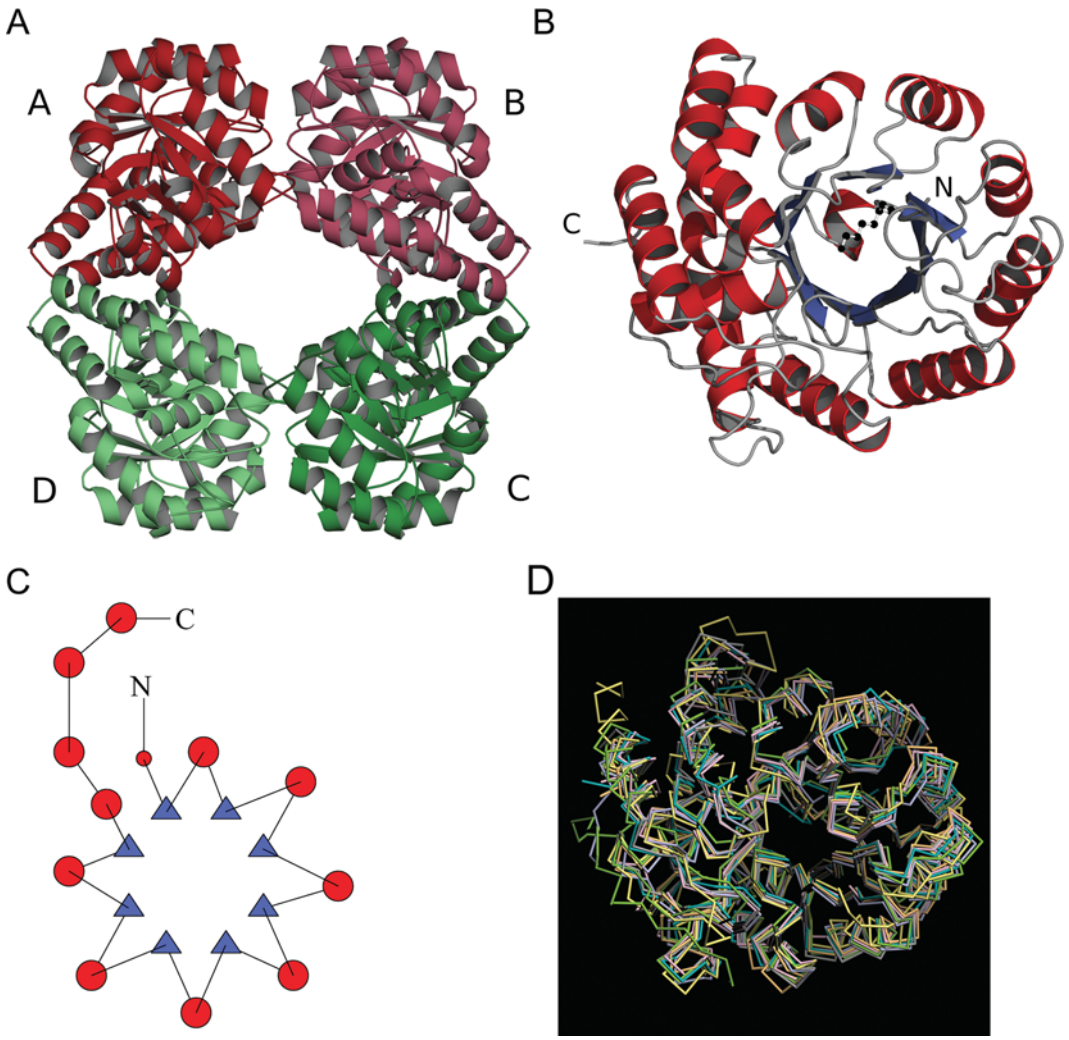


Figure 1 The oligomeric and monomeric structure of the *Mtb*-DHDPS

(**A**) The two monomers forming the tight dimer AB are coloured red and light red, and the two monomers forming the second tight dimer CD are coloured green and light green respectively. Monomer structure (**B**) and topology plot (**C**) of *Mtb*-DHDPS. (**D**) Superposition of DHDPS structures on to the A-subunit of *Mtb*-DHDPS (1XXX, cyan). R.m.s.d. (root mean square deviation) values (in Å) and number of superposed residues (out of 296), are also given in parentheses: *E. coli* (1YXC, purple, 1.37, 285); *T. maritima* (1O5K, orange, 1.32, 286); *N. sylvestris* (structural data from R. Huber, green, 1.53, 276); *B. anthracis* (1XKY, grey, 1.17, 287); and *A. tumefaciens* (2HMC, yellow, 1.75, 272).

described as a dimer of dimers, with the two monomers A and B (and C and D, Figure 1A) tightly bound to each other to form the tight dimer, and weaker interactions between the AB and CD dimeric units.

The final refined *Mtb*-DHDPS model contains 2364 residues in total, 1587 water molecules, eight DTT molecules covalently bound to Cys^{248}, eight Mg^{2+} and eight Cl^{-} ions. The model includes all amino acid residues, with the exception of the first four or five amino acids at the N-terminus, which are not visible in the electron density. Of all residues, 90.4 % are in the most favoured regions of the Ramachandran plot [41] and 9.2 % of residues are in additionally allowed regions. Only Tyr^{117} lies in the forbidden region of the Ramachandran plot, with the exact values being $\varphi = 74.9°$ and $\psi = -52.5°$ for Tyr^{117} in chain A. The unusual conformation of Tyr^{117} is well supported by the electron density in each of the monomers. The same observation has been made previously for the corresponding residue in all other DHDPS enzymes of known structure. Tyr^{117} lies at the tight-dimer interface of two monomers and takes part in formation of the active site of the neighbouring monomer.

A secondary structure matching search against all structures deposited in the PDB using subunit A of *Mtb*-DHDPS yielded a total of 21 hits above a Q-score threshold of 0.32. After removing redundancies, 12 entries were left (Q-score = 0.57–0.78) (see Supplementary Table 1 at http://www.BiochemJ.org/bj/411/bj4110351add.htm). The most similar structure was found to be DHDPS from *B. anthracis* followed by DHDPS from *T. maritima*. At the time of solving the structure of *Mtb*-DHDPS, only the latter was available in the PDB. This in retrospect validates the choice of the search model for molecular replacement, which was based solely on amino acid sequence comparison. An interesting observation is the apparent relationship of DHDPS to *N*-acetylneuraminate lyases, which has been noted previously [42].

Active site and metal-binding sites

The active site is located at the centre of each monomer, facing the central cavity of the tetramer. It is situated in a pocket at the C-terminal side of the $(\beta/\alpha)_8$-barrel, as in all known $(\beta/\alpha)_8$-enzymes. The active site centres about Lys^{171}, the amino acid that

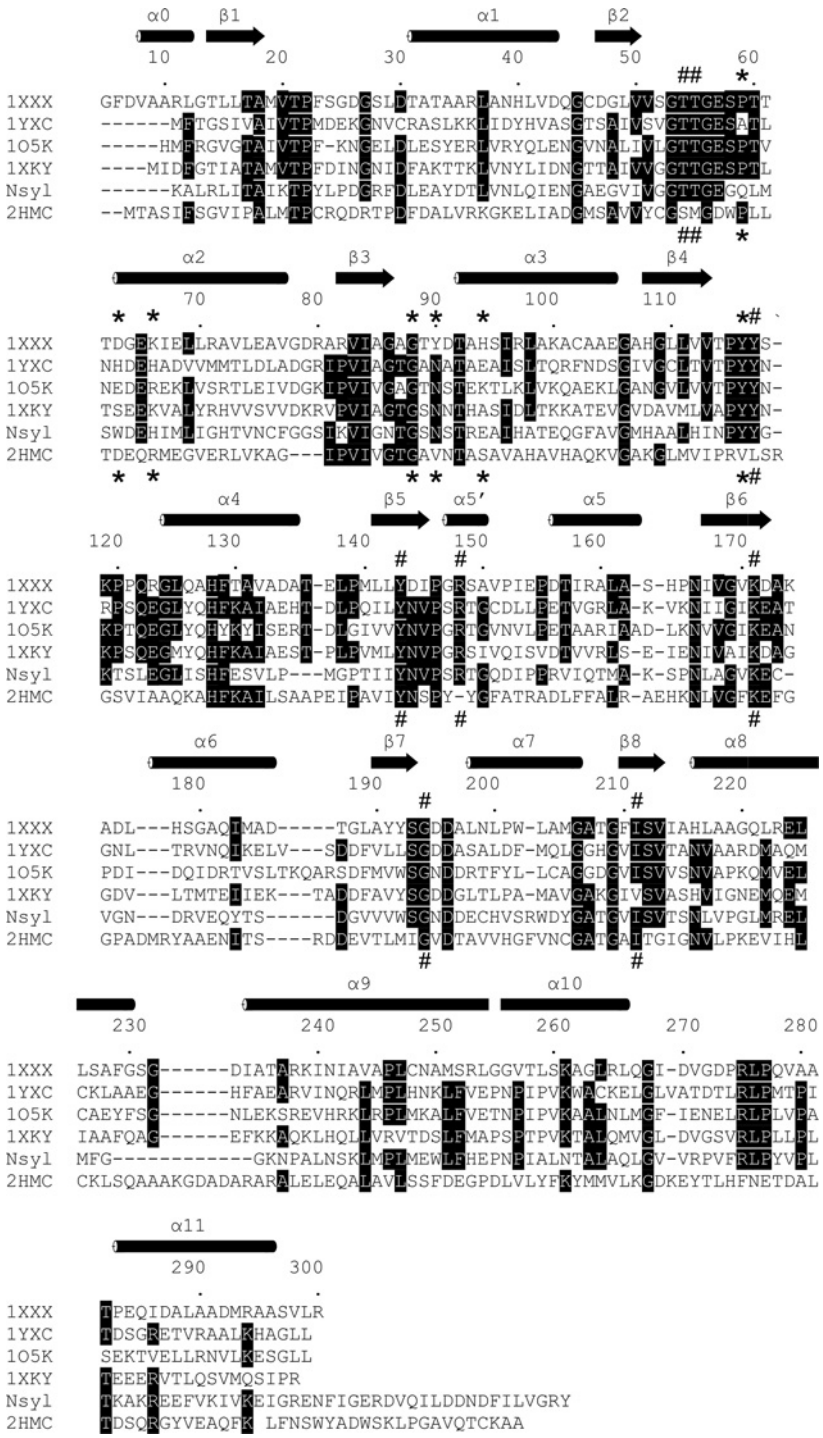


Figure 2 Multiple sequence alignment

Multiple sequence alignment of the DHDPS sequences from *M. tuberculosis* (1XXX), *E. coli* (1YXC), *T. maritima* (1O5K), *B. anthracis* (1XKY), *N. sylvestris* (Nsyl) and *A. tumefaciens* (2HMC). The alignment was derived from the three-dimensional alignments of the DHDPS structures, which were carried out using the programs ALIGN [29] and STAMP [30]. The secondary-structure elements (α-helices and β-strands) observed in the structure of *Mtb*-DHDPS are indicated and labelled. Amino acid residues in black boxes are conserved in four or more of the six DHDPS enzymes listed. Hash marks (#) indicate those residues involved in catalysis, and an asterisk (*) indicates those residues that are involved in (*S*)-lysine binding in the *E. coli*, *T. maritima* and *N. sylvestris* enzymes.

forms a Schiff base with the first substrate, pyruvate. In the *E. coli* enzyme, the equivalent amino acid is Lys[161], as previously identified via tryptic digest studies [7]. Three amino acid residues form the conserved catalytic triad: Tyr[143], Thr[54], and Tyr[117], whereby Tyr[117] is contributed from the adjacent monomer across the tight-dimer interface. Both the identity, as well as the relative spatial orientation of these functional groups, is conserved among characterized DHDPS enzymes, but not conserved in the putative *A. tumefaciens* DHDPS enzyme (Swiss-Prot accession number Q8U6Y1), where the corresponding residues are Tyr[136], Ser[48] and Leu[108] (Figure 2). As mentioned above, the exact function of the *A. tumefaciens* enzyme remains to be established. A detailed

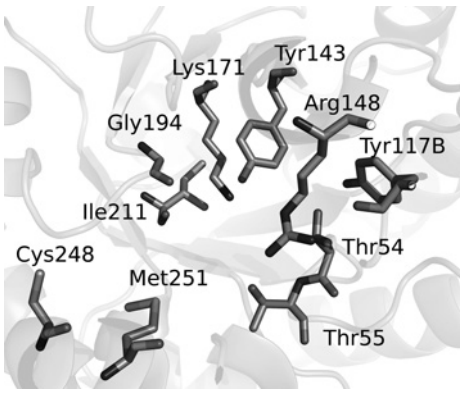


Figure 3 Survey of the *Mtb*-DHDPS active site

The active site, defined by the position of Lys^{171}, is found in a cleft within the $(\alpha/\beta)_8$-barrel (secondary structure is transparent). Tyr^{117} from the adjacent monomer reaches through the interface (AB from Figure 1) to interact with Thr^{54}.

examination of the *Mtb*-DHDPS active site, however, shows the distance between Tyr^{143}-OH and Thr^{54}-OH is somewhat increased (4.7 Å) compared with *E. coli* DHDPS (2.7 Å), and the geometry is such that hydrogen-bonding is unlikely (Figure 3). In addition, the active site also contains a methionine residue (Met^{251}) and a cysteine residue (Cys^{248}), which, in *Mtb*-DHDPS, binds a DTT molecule. These residues sit opposite Lys^{171}, lining the active-site cavity (Figure 3), and could be exploited in rational inhibitor design. We have noted recently the potential for a species-specific inhibitor designed to target DHDPS [43]. Whether these motifs are utilized in *Mtb*-DHDPS catalysis is under investigation. The mechanism of the DHDPS-catalysed reaction has been studied in detail for the *E. coli* enzyme by X-ray crystallography, NMR [10] and site-directed mutagenesis [12]. After Schiff base formation between pyruvate and the active-site lysine residue, (*S*)-ASA binds to the stable substituted enzyme form. The product heterocycle (HTPA) is then released and undergoes dehydration and the enzyme is regenerated [10].

An Mg^{2+} ion is co-ordinated by Ala^{162}-O, His^{164}-O and Ile^{167}-O and three water molecules. In the *E. coli* (PDB codes 1DHP and 1YXC [13,14]) and *B. anthracis* (PDB entry 1XKY [16]) DHDPS structures, a K^+ ion is found in the same position as the Mg^{2+} ion in the *Mtb*-DHDPS structure. The function of this metal-binding site is unclear, especially since it is located far from the enzyme active site. We note that, via the coupled assay, the activity of the enzyme is unaffected by Mg^{2+} ions (< 20 mM). It is likely that the Mg^{2+} ions seen in the structure are an artefact of the crystallization conditions. A Cl^- ion, co-ordinated by Lys^{174}-NZ, Ser^{179}-OG and two water molecules, was also observed in *Mtb*-DHDPS. In contrast with the model of the *E. coli* enzyme (PDB code 1YXC [13]), where a Cl^- ion is found in the active site, co-ordinated by the catalytic Lys^{161}, Thr^{44} and two water molecules [13], the Cl^--binding site in *Mtb*-DHDPS is distal to the active site, leaving its functional significance unclear.

Dimer–dimer interface of *Mtb*-DHDPS

In order to confirm the biologically significant unit of the enzyme, all the potential interfaces were examined using the EBI PISA server [32], and the Protein–Protein Interaction server. Within the tetrameric unit shown in Figure 1(A), each of the two tight-dimer interfaces (AB and CD) bury approx. 1470 Å^2 per monomer, which corresponds to roughly 13 % of the surface area of the monomer, whereas the interfaces between monomers A and D, and monomers B and C, bury a surface area of approx. 860 Å^2 per monomer (see Supplementary Table 2 at http://www.BiochemJ.org/bj/411/bj4110351add.htm). Polar residues represent 40 % of the tight interface area, and 30 % of the AD and BC interfaces. For the *E. coli* enzyme, each of the two tight-dimer interfaces (AB and CD) bury approx. 1300 Å^2 per monomer (roughly 11 % of the surface area of the monomer), whereas the weaker dimer interfaces AC and BD bury a surface area of approx. 500 Å^2. The polar residues represent 35 % of the tight-dimer interface area, and 44 % of the weak AC and BD interfaces. The *T. maritima* DHDPS structure buries an interface-accessible area similar to that of the *Mtb*-DHDPS structure (see Supplementary Table 2). Figure 4(A) shows a schematic diagram of the *E. coli* DHDPS and *Mtb*-DHDPS dimer–dimer interface. This comparison suggests that the oligomeric structure of *Mtb*-DHDPS may be somewhat more stable than the structure of *E. coli* DHDPS.

In order to corroborate the apparent oligomeric state seen in the crystal structure, further study of *Mtb*-DHDPS in solution was performed via analytical ultracentrifugation. Sedimentation velocity experiments, fitted to a continuous size-distribution model [38], confirmed that *Mtb*-DHDPS was a single species in solution with a molecular mass of 115 kDa, taken from the ordinate maximum of the peak observed in the $c(M)$ distribution (Figure 4C). Based on $c(s)$ distribution analysis (results not shown), the tetrameric species has a sedimentation coefficient ($S^0_{20,w}$) of 6.7 S. We note that, at a concentration of 0.055 mg · ml^{-1}, *Mtb*-DHDPS shows no dimer in solution. This is in contrast with the *E. coli* enzyme, which shows a significant proportion of dimer at a similar concentration (0.05 mg · ml^{-1}) [44], suggesting that the dimer–dimer interface for *Mtb*-DHDPS is indeed stronger than in the *E. coli* enzyme. Using eqn (1), the spatial arrangement of the *Mtb*-DHDPS tetramer in aqueous solution ($S^0_{20,w} = 6.7$ S) is intermediate between square planar (calculated $S^0_{20,w} = 6.540$ S, $F = 0.926$) and tetrahedral (calculated $S^0_{20,w} = 6.903$ S, $F = 0.977$), which correlates well with the particular D_2-symmetric shape shown in the X-ray crystal model (Figure 1). The homotetrameric structure of *Mtb*-DHDPS in solution has been corroborated further by gel filtration and blue native PAGE (results not shown).

Enzyme kinetics

The coupled enzyme assay was used to characterize the kinetic properties of *Mtb*-DHDPS. Initially, the stability and pH optimum for activity was assessed. For enzymatic activity of *Mtb*-DHDPS, the optimum pH was determined to be 8.25 at 30 °C in Hepes buffer (Figure 5A). Additionally, a steady decrease in activity was seen when the ionic strength was increased above 0.1 M (Figure 5B). In a buffer of 20 mM Hepes (pH 8.25) at 30 °C (ionic strength 0.15 M), *Mtb*-DHDPS showed a greater thermal stability than *E. coli* DHDPS, with an apparent T_m (melting temperature) of ~82 °C compared with ~57 °C for the *E. coli* enzyme (Figure 5C). Further experiments showed that *Mtb*-DHDPS maintained its activity for 40 min when incubated at 70 °C, whereas *E. coli* DHDPS showed degradation within the first few minutes (results not shown). This result was supported by following protein unfolding monitored by fluorescence emission (Figure 5D). The thermal stability of the *Mtb*-DHDPS is probably in part due to the greater number of intersubunit contacts as compared with the *E. coli* enzyme (Figure 4A).

A full matrix of initial rates was determined with both substrates varied, and these data were fitted with different kinetic models: ternary complex, Ping Pong and Ping Pong with substrate inhibition. The Ping Pong model provided the best fit (Figure 6), and yielded K_M constants for pyruvate (0.17 ± 0.01 mM) and (*S*)-ASA

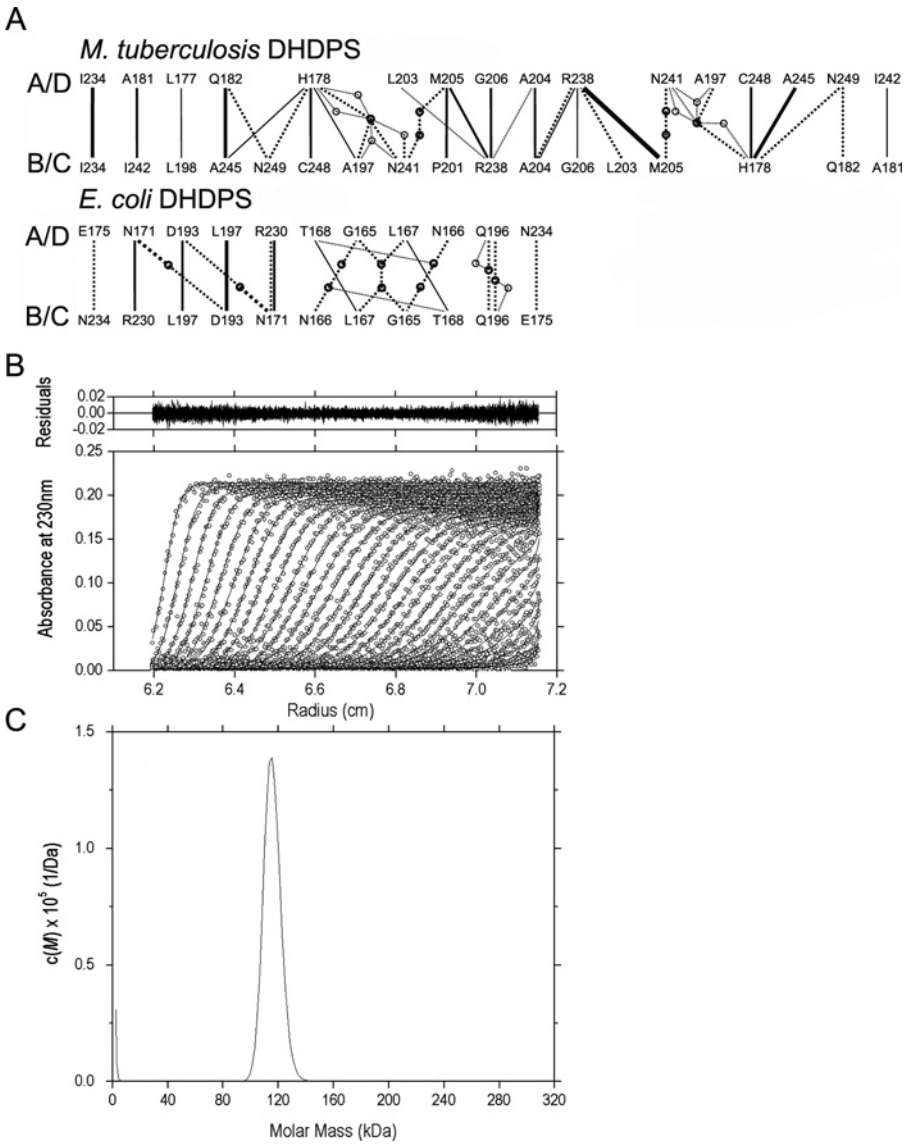


Figure 4 Analysis of the dimer–dimer interface and sedimentation velocity of *Mtb*-DHDPS

(**A**) Analysis of the dimer–dimer interface of tetrameric *Mtb*-DHDPS (1XXX) and *E. coli* DHDPS (1YXC). Interfaces were probed using [Java]Protein Dossier [51] and manual inspection. Dotted lines indicate hydrogen-bonding, continuous lines indicate hydrophobic interactions, and circles represent water. The thicker lines indicate that the interactions are in both symmetrical interfaces (i.e. interfaces AD and BC of Figure 1) of the tetramer. Amino acids are identified using single-letter codes. (**B**) Sedimentation velocity absorbance data plotted as a function of radial position from the axis of rotation (cm) for *Mtb*-DHDPS at a concentration of 0.055 mg · ml^{-1} in 20 mM Tris/HCl and 150 mM NaCl, at pH 8.0. The raw data are presented as open symbols (○) plotted at time intervals of 6 min overlaid with the non-linear least squares best-fit (solid line) to a continuous size distribution model [$c(M)$] [38]. The residuals (top) for the $c(M)$ distribution best-fit are plotted as a function of radial position (cm) from the axis of rotation. (**C**) The $c(M)$ distribution is plotted as a function of molar mass (kDa) for *Mtb*-DHDPS. The fit was obtained using a resolution of 200 species between M_{min} of 2.5 kDa and M_{max} of 350 kDa with $\bar{\nu} = 0.7402$ ml · g^{-1}, $\rho = 1.00\,499$ g · ml^{-1}, $\eta = 1.0214$ cP, and $f/f_0 = 1.21902$. The r.m.s.d. (root mean square deviation) and Z test for the fit were 0.003895 and 3.49 respectively.

(0.43 ± 0.02 mM). Interestingly, a 6-fold increase in the V_{max} (4.42 ± 0.08 μmol · s^{-1} · mg^{-1}) was observed over *E. coli* DHDPS, which may reflect the increased importance of DAP in the peptidoglycan layer of mycobacteria [45]. That the ordered Ping Pong kinetic mechanism is conserved in *Mtb*-DHDPS, is consistent with the ordered kinetic mechanism proposed for other DHDPS enzymes. We note that the Ping Pong mechanism is symmetrical, but speculate that, in the case of *Mtb*-DHDPS, pyruvate is the first substrate to bind, followed by (*S*)-ASA. Whether the catalytic triad, as proposed for the *E. coli* enzyme, operates in this enzyme remains to be elucidated; as we have observed above, this motif is conserved, yet its geometry probably precludes function without structural changes occurring, perhaps during catalysis. The structure of *Mtb*-DHDPS complexed with various inhibitors or substrates will be helpful in this regard.

Previous studies suggest that the real power of an enzyme lies not only in the mechanistically important residues that form the catalytic unit, but also in a combination of the very 'local' structural features of the catalytic unit and more 'global' features, such as the dynamics of the structure and the overall microenvironment of the active site [46]. Despite the high degree of overall structure similarity (both in the solid state and in solution) and active-site conservation, *Mtb*-DHDPS shares little sequence identity with other DHDPS enzymes, and our data reveal important differences in their structural, biophysical and biochemical properties.

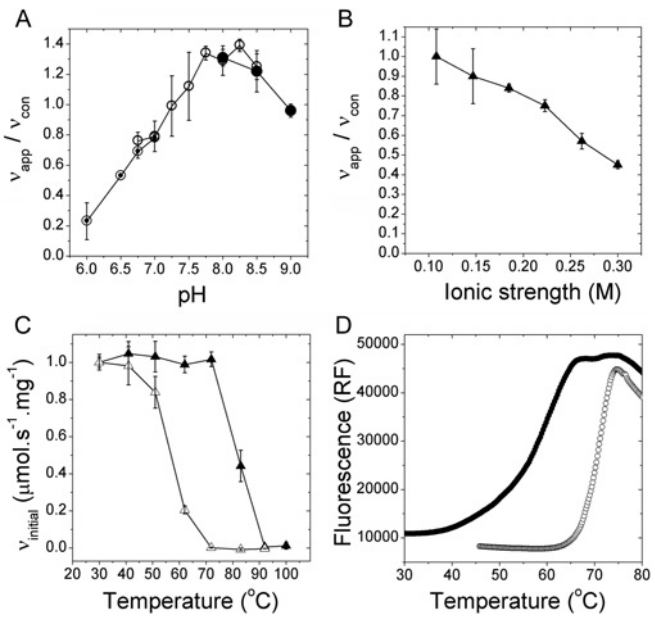


Figure 5 Effect of pH, ionic strength and temperature on catalytic activity

The effect of varying (**A**) pH with different buffers (⊙ Mes, ○ Hepes and ● Bicine), and (**B**) ionic strength on enzyme activity was measured using the coupled assay. The data in (**B**) are normalized to the highest measured rate (v_{con}) of enzyme activity. (**C**) Enzyme activity determined using the coupled assay after pre-treatment using a 5 min incubation at increasing temperature. The thermal stability of *Mtb*-DHDPS (▲) was compared with *E. coli* DHDPS (△) using the apparent initial rate (v_{app}) divided by rate of enzyme activity without pre-treatment (v_{con}). Results are means ± S.D. of triplicate measurements. (**D**) The melting temperature as determined by a fluorescence melt for *Mtb*-DHDPS (○) and *E. coli* DHDPS (●) was comparable with that determined using activity assays.

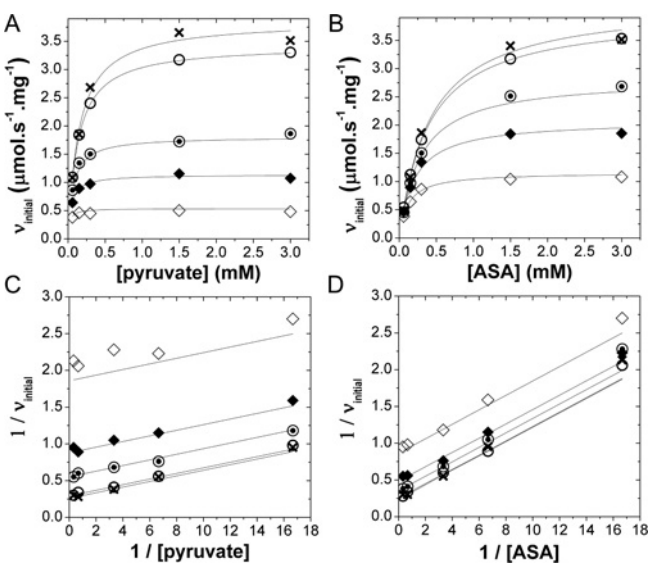


Figure 6 Kinetic analysis of *Mtb*-DHDPS

The initial velocity was measured using the coupled assay over various concentrations of each substrate, (*S*)-ASA and pyruvate (× 3.0 mM, ○ 1.5 mM, ⊙ 0.30 mM, ◆ 0.15 mM and ◇ 0.06 mM). Each point was measured at least in triplicate and the data were fitted with the Ping Pong model (R^2 of 0.9945) using the software program Enzfitter (**A** and **B**). (**C**, **D**) Lineweaver–Burk transformations reflect the trend predicted for the Ping Pong model.

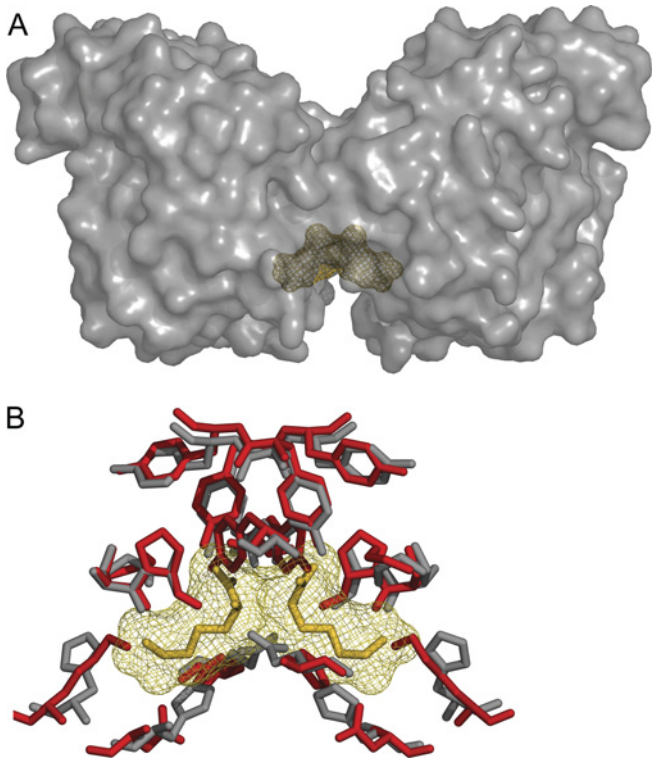


Figure 7 Comparison of the (*S*)-lysine-binding site in *E. coli* DHDPS and cleft in *Mtb*-DHDPS

(**A**) Surface representation of the tight dimer showing the (*S*)-lysine-binding site of the *E. coli* DHDPS (1YXD) with the bound (*S*)-lysine in yellow. (**B**) Superposition of the corresponding *Mtb*-DHDPS (1XXX) residues (in red) on to the (*S*)-lysine-binding residues (grey) of *E. coli* DHDPS (1YXD). The bound (*S*)-lysine residues (from 1YXD) are shown in yellow.

Mtb-DHDPS is not feedback-regulated by (*S*)-lysine

Except at very high concentrations (50 mM and above), (*S*)-lysine did not have any effect on the activity of *Mtb*-DHDPS. The IC_{50} was determined to be 250 mM, which greatly exceeds that which could be reasonably expected in a cell. Thus, in contrast with plant and Gram-negative DHDPS enzymes [8,35,47–49], *Mtb*-DHDPS can be considered insensitive to inhibition by (*S*)-lysine. This observation is explained nicely by structural superposition of the two enzymes and examination of the (*S*)-lysine-binding site observed in *E. coli* (Figure 7). Most residues identified as important for the (*S*)-lysine allosteric binding site are not conserved in *Mtb*-DHDPS (Table 2) [9,10,13]. *Mtb*-DHDPS is the only DHDPS enzyme with a known structure in which the asparagine residue of the (*S*)-lysine-binding site is not conserved, but replaced by a tyrosine residue. Additionally, the (*S*)-lysine-binding cavity in the *M. tuberculosis* enzyme is shallow compared with the equivalent site of the *E. coli* or *N. sylvestris* enzymes, leaving little space for (*S*)-lysine to bind. Interestingly, other aspartate family amino acids (DAP, threonine and methionine) showed no significant inhibition of *Mtb*-DHDPS activity (Table 3).

The regulation of the (*S*)-lysine biosynthetic pathway (if it is indeed regulated in *M. tuberculosis*) is maintained via an alternative strategy. In *E. coli*, the DAP pathway is likely to be controlled at both the aspartate kinase and DHDPS catalytic step. Moreover, it has been suggested that *E. coli* DHDPS expression is regulated by the level of DAP [50]; whether such an approach is adopted by *M. tuberculosis* remains to be elucidated.

Table 2 Allosteric inhibition site residues

The (*S*)-lysine-binding residues are identified in *E. coli* DHDPS (second column), followed by an alignment of corresponding residues in the other known structures. *N. sylvestris* DHDPS co-ordinates were kindly provided by R. Huber.

M. tuberculosis (1XXX)	*E. coli* (1YXD)	*T. maritima* (1O5K)	*N. sylvestris*	*B. anthracis* (1XKY)
Ser^{58}	Ser	Ser	Gly	Ser
Pro^{59}	Ala	Pro	Gln	Pro
Asp^{63}	His	Glu	Trp	Ser
Lys^{66}	His	Arg	His	Lys
Gly^{88}	Gly	Gly	Gly	Gly
Tyr^{90}	Asn	Asn	Asn	Asn
His^{94}	Glu	Lys	Glu	Ala
Tyr^{116}	Tyr	Tyr	Tyr	Tyr

Table 3 Potential inhibitors of *Mtb*-DHDPS

The activity of *Mtb*-DHDPS in the presence of aspartate family amino acids. Each substrate concentration was held at 0.3 mM. Relative activity is relative to that in the absence of inhibitor.

Amino acid	Concentration (mM)	Relative activity (%)
meso-DAP	1	99
	10	87
	20	57
(*S*)-threonine	1	104
	10	86
	20	74
(*S*)-methionine	1	100
	10	99
	20	96

Accordingly, the most valid approach for generating inhibitors and novel antibiotics targeting DHDPS from *M. tuberculosis* should focus on the highly conserved active-site geometry of the enzyme, rather than the vestigial allosteric cleft (Figure 7).

In conclusion, the increasing prevalence of tuberculosis cases, and especially that of antibiotic resistance, necessitates the development of novel antibiotics. Pivotal to such endeavours is an extensive understanding of the proposed antibiotic targets. In the present paper, we describe a structural and biochemical study of an important antibiotic target from *M. tuberculosis*. We have found that the active site of *Mtb*-DHDPS is generally similar to that of *E. coli* DHDPS. However, we note that the proton-relay residues are such that the hydrogen-bonding network may be disrupted. Given that the residues responsible for substrate binding are generally conserved, lead compounds targeted to the *E. coli* active site are likely to also be effective against the *M. tuberculosis* enzyme. Additionally, the presence of a cysteine residue and a methionine residue within the active site may be exploited to develop inhibitors tailored to *Mtb*-DHDPS. The *M. tuberculosis* enzyme, unlike *E. coli* DHDPS, is not allosterically affected by (*S*)-lysine, which may reflect the requirement for DAP in the bacterial cell wall. Continued investigation into the catalytic mechanisms, the consequences of this enzyme's oligomeric state and the DAP/(*S*)-lysine biosynthetic pathway in *M. tuberculosis* are necessary to explain the regulatory mechanism of this essential pathway in bacteria.

We thank Dr Jeanne Perry (UCLA, Los Angeles, CA, U.S.A.) for providing genomic *Mtb*-DNA, Dr Arie Geerlof (EMBL Hamburg) for help with the solubility screens, Dr Santosh Panjikar (EMBL Hamburg) for help in solving the *Mtb*-DHDPS structure by molecular replacement, Dr Robert Huber (Cardiff University, Cardiff, Wales, U.K.) for providing the structural co-ordinates of *N. sylvestris* and *E. coli* DHDPS. We thank Dr George Lorimer (E.I. Dupont De Nemours and Co., Wilmington, DE, U.S.A.) for generously providing the pGroESL plasmid, and Jackie Healy for unassuming and resourceful technical support. G.K. was funded by the TB consortium (http://www.doe-mbi.ucla.edu/TB) for postdoctoral exchange grants and the X-Mtb consortium (http://www.xmtb.org) for funding through BMBF (Bundesministerium für Bildung und Forschung)/PTJ (Projektträger Jülich) grant number BIO/0312992A. G.L.E., J.A.G., M.A.P., M.D.W.G. and R.C.J.D. were supported in part by the Royal Society of New Zealand Marsden Fund (contract UOC303), S.R.A.D. acknowledges financial support from FRST (Foundation for Research, Science and Technology), and M.A.P., J.A.G. and R.C.J.D. were supported in part by the DTRA (Defense Threat Reduction Agency) (project W911NF-07-1-0105).

REFERENCES

1 Bloom, B. R. and Murray, C. J. (1992) Tuberculosis: commentary on a reemergent killer. Science **257**, 1055–1064
2 Meya, D. B. and McAdam, K. P. (2007) The TB pandemic: an old problem seeking new solutions. J. Intern. Med. **261**, 309–329
3 World Health Organization (2007) Global Tuberculosis Control: Surveillance, Planning, Financing, World Health Organization, Geneva, http://www.who.int/tb/publications/global_report/2007/pdf/full.pdf
4 Pavelka, M. S. and Jacobs, W. R. (1996) Biosynthesis of diaminopimelate, the precursor of lysine and a component of peptidoglycan, is an essential function of *Mycobacterium smegmatis*. J. Bacteriol. **178**, 6496–6507
5 Cox, R. J., Sutherland, A. and Vederas, J. C. (2000) Bacterial diaminopimelate metabolism as a target for antibiotic design. Bioorg. Med. Chem. Lett. **8**, 843–871
6 Hutton, C. A., Southwood, T. J. and Turner, J. J. (2003) Inhibitors of lysine biosynthesis as antibacterial agents. Mini Rev. Med. Chem. **3**, 115–127
7 Laber, B., Gomis-Ruth, F. X., Romao, M. J. and Huber, R. (1992) *Escherichia coli* dihydrodipicolinate synthase: identification of the active site and crystallization. Biochem. J. **288**, 691–695
8 Shedlarski, J. G. and Gilvarg, C. (1970) The pyruvate–aspartic semialdehyde condensing enzyme of *Escherichia coli*. J. Biol. Chem. **245**, 1362–1373
9 Blickling, S., Beisel, H. G., Bozic, D., Knablein, J., Laber, B. and Huber, R. (1997) Structure of dihydrodipicolinate synthase of *Nicotiana sylvestris* reveals novel quaternary structure. J. Mol. Biol. **274**, 608–621
10 Blickling, S., Renner, C., Laber, B., Pohlenz, H. D., Holak, T. A. and Huber, R. (1997) Reaction mechanism of *Escherichia coli* dihydrodipicolinate synthase investigated by X-ray crystallography and NMR spectroscopy. Biochemistry **36**, 24–33
11 Karsten, W. E. (1997) Dihydrodipicolinate synthase from *Escherichia coli*: pH dependent changes in the kinetic mechanism and kinetic mechanism of allosteric inhibition by L-lysine. Biochemistry **36**, 1730–1739
12 Dobson, R. C. J., Valegard, K. and Gerrard, J. A. (2004) The crystal structure of three site-directed mutants of *Escherichia coli* dihydrodipicolinate synthase: further evidence for a catalytic triad. J. Mol. Biol. **338**, 329–339
13 Dobson, R. C. J., Griffin, M. D. W., Jameson, G. B. and Gerrard, J. A. (2005) The crystal structures of native and (*S*)-lysine-bound dihydrodipicolinate synthase from *Escherichia coli* with improved resolution show new features of biological significance. Acta Crystallogr. Sect. D Biol. Crystallogr. **61**, 1116–1124
14 Mirwaldt, C., Korndorfer, I. and Huber, R. (1995) The crystal structure of dihydrodipicolinate synthase from *Escherichia coli* at 2.5 Å resolution. J. Mol. Biol. **246**, 227–239
15 Pearce, F. G., Perugini, M. A., McKerchar, H. J. and Gerrard, J. A. (2006) Dihydrodipicolinate synthase from *Thermotoga maritima*. Biochem. J. **400**, 359–366
16 Blagova, E., Levdikov, V., Milioti, N., Fogg, M. J., Kalliomaa, A. K., Brannigan, J. A., Wilson, K. S. and Wilkinson, A. J. (2006) Crystal structure of dihydrodipicolinate synthase (BA3935) from *Bacillus anthracis* at 1.94 Å resolution. Proteins **62**, 297–301
17 Dobson, R. C. J., Devenish, S. R., Turner, L. A., Clifford, V. R., Pearce, F. G., Jameson, G. B. and Gerrard, J. A. (2005) Role of arginine 138 in the catalysis and regulation of *Escherichia coli* dihydrodipicolinate synthase. Biochemistry **44**, 13007–13013
18 Zor, T. and Selinger, Z. (1996) Linearization of the Bradford protein assay increases its sensitivity: theoretical and experimental studies. Anal. Biochem. **236**, 302–308
19 Kefala, G. and Weiss, M. S. (2006) Cloning, expression, purification, crystallization and preliminary X-ray diffraction analysis of DapA (Rv2753c) from *Mycobacterium tuberculosis*. Acta Crystallograph. Sect. F Struct. Biol. Crystal. Commun. **62**, 1116–1119
20 Goloubinoff, P., Gatenby, A. A. and Lorimer, G. H. (1989) GroE heat-shock proteins promote assembly of foreign prokaryotic ribulose bisphosphate carboxylase oligomers in *Escherichia coli*. Nature **337**, 44–47
21 Yugari, Y. and Gilvarg, C. (1965) The condensation step in diaminopimelate synthesis. J. Biol. Chem. **240**, 4710–4716

22 Schagger, H., Cramer, W. A. and von Jagow, G. (1994) Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. Anal. Biochem. **217**, 220–230
23 Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S. et al. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. Sect. D Biol. Crystallogr. **54**, 905–921
24 Perrakis, A., Morris, R. and Lamzin, V. S. (1999) Automated protein model building combined with iterative structure refinement. Nat. Struct. Biol. **6**, 458–463
25 Jones, T. A., Zou, J. Y., Cowan, S. W. and Kjeldgaard, M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. Sect. A Found. Crystallogr. **47**, 110–119
26 Murshudov, G. N., Vagin, A. A. and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. Sect. D Biol. Crystallogr. **53**, 240–255
27 Laskowski, R. A., Macarthur, M. W., Moss, D. S. and Thornton, J. M. (1993) Procheck: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. **26**, 283–291
28 Reference deleted
29 Cohen, G. E. (1997) ALIGN: a program to superimpose protein coordinates, accounting for insertions and deletions. J. Appl. Crystallogr. **30**, 1162–1164
30 Russell, R. B. and Barton, G. J. (1992) Multiple protein sequence alignment from tertiary structure comparison: assignment of global and residue confidence levels. Proteins **14**, 309–323
31 Kabsch, W. (1976) Solution for best rotation to relate 2 sets of vectors. Acta Crystallogr. Sect. A Cryst. Phys. Diffr. Theor. Gen. Crystallogr. **32**, 922–923
32 Krissinel, E. and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. **372**, 774–797
33 Roberts, S. J., Morris, J. C., Dobson, R. C. J. and Gerrard, J. A. (2003) The preparation of (*S*)-aspartate semi-aldehyde appropriate for use in biochemical studies. Bioorg. Med. Chem. Lett. **13**, 265–267
34 Dobson, R. C. J., Gerrard, J. A. and Pearce, F. G. (2004) Dihydrodipicolinate synthase is not inhibited by its substrate, (*S*)-aspartate β-semialdehyde. Biochem. J. **377**, 757–762
35 Dobson, R. C. J., Griffin, M. D. W., Roberts, S. J. and Gerrard, J. A. (2004) Dihydrodipicolinate synthase (DHDPS) from *Escherichia coli* displays partial mixed inhibition with respect to its first substrate, pyruvate. Biochimie **86**, 311–315
36 Selwyn, M. J. (1965) A simple test for inactivation of an enzyme during assay. Biochim. Biophys. Acta **105**, 193–195
37 Ericsson, U. B., Hallberg, B. M., Detitta, G. T., Dekker, N. and Nordlund, P. (2006) Thermofluor-based high-throughput stability optimization of proteins for structural studies. Anal. Biochem. **357**, 289–298
38 Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. Biophys. J. **78**, 1606–1619
39 Laue, T. M., Shah, D. B., Ridgeway, T. M. and Pelletier, S. L. (1992) Computer-aided interpretation of analytical sedimentation data for proteins. In Analytical Ultracentrifugation in Biochemistry and Polymer Science (Harding, S. E., Horton, J. C. and Rowe, A. J., eds.), pp. 90–125, The Royal Society of Chemistry, Cambridge
40 Teller, D. C., Swanson, E. and de Haen, C. (1979) The translational friction coefficient of proteins. Methods Enzymol. **61**, 103–124
41 Ramakrishnan, C. and Ramachandran, G. N. (1965) Stereochemical criteria for polypeptide and protein chain conformations. II. Allowed conformations for a pair of peptide units. Biophys. J. **5**, 909–933
42 Lawrence, M. C., Barbosa, J. A. R. G., Smith, B. J., Hall, N. E., Pilling, P. A., Ooi, H. C. and Marcuccio, S. M. (1997) Structure and mechanism of a sub-family of enzymes related to *N*-acetylneuraminate lyase. J. Mol. Biol. **266**, 381–399
43 Mitsakos, V., Dobson, R. C. J., Pearce, F. G., Devenish, S. R., Evans, G. L., Burgess, B. R., Perugini, M. A., Gerrard, J. A. and Hutton, C. A. (2008) Inhibiting dihydrodipicolinate synthase across species: towards specificity for pathogens? Bioorg. Med. Chem. Lett. **18**, 842–844
44 Perugini, M. A., Griffin, M. D., Smith, B. J., Webb, L. E., Davis, A. J., Handman, E. and Gerrard, J. A. (2005) Insight into the self-association of key enzymes from pathogenic species. Eur. Biophys. J. **34**, 469–476
45 Wietzerbin, J., Das, B. C., Petit, J. F., Lederer, E., Leyh-Bouille, M. and Ghuysen, J. M. (1974) Occurrence of D-alanyl-(D)-*meso*-diaminopimelic acid and *meso*-diaminopimelyl-*meso*-diaminopimelic acid interpeptide linkages in the peptidoglycan of *Mycobacteria*. Biochemistry **13**, 3471–3476
46 Gutteridge, A. and Thornton, J. M. (2005) Understanding nature's catalytic toolkit. Trends Biochem. Sci. **30**, 622–629
47 Frisch, D. A., Gengenbach, B. G., Tommey, A. M., Sellner, J. M., Somers, D. A. and Myers, D. E. (1991) Isolation and characterization of dihydrodipicolinate synthase from maize. Plant Physiol. **96**, 444–452
48 Mazelis, M., Whatley, F. R. and Whatley, J. (1977) The enzymology of lysine biosynthesis in higher plants: the occurrence, characterization and some regulatory properties of dihydrodipicolinate synthase. FEBS Lett. **84**, 236–240
49 Wallsgrove, R. M. and Mazelis, M. (1981) Spinach leaf dihydrodipicolinate synthase: partial purification and characterization. Phytochemistry **20**, 2651–2655
50 Acord, J. and Masters, M. (2004) Expression from the *Escherichia coli* dapA promoter is regulated by intracellular levels of diaminopimelic acid. FEMS Microbiol. Lett. **235**, 131–137
51 Neshich, G., Rocchia, W., Mancini, A. L., Yamagishi, M. E., Kuser, P. R., Fileto, R., Baudet, C., Pinto, I. P., Montagner, A. J., Palandrani, J. F. et al. (2004) JavaProtein Dossier: a novel web-based data visualization tool for comprehensive analysis of protein structure. Nucleic Acids Res. **32**, W595–W601

Received 4 October 2007/27 November 2007; accepted 7 December 2007
Published as BJ Immediate Publication 7 December 2007, doi:10.1042/BJ20071360

Appendix K

Journal publication

Inhibiting dihydrodipicolinate synthase across species: towards specificity for pathogens?

Mitsakos, V., Dobson, R. C., Pearce, F. G., Devenish, S. R., Evans, G. L., Burgess, B. R., Perugini, M. A., Gerrard, J. A. & Hutton, C. A.

Available online at www.sciencedirect.com
ScienceDirect
Bioorganic & Medicinal Chemistry Letters 18 (2008) 842–844
Bioorganic & Medicinal Chemistry Letters

Inhibiting dihydrodipicolinate synthase across species: Towards specificity for pathogens?

Voula Mitsakos,[a,b] Renwick C. J. Dobson,[b,c,d] F. Grant Pearce,[d] Sean R. Devenish,[d] Genevieve L. Evans,[d] Benjamin R. Burgess,[b,c] Matthew A. Perugini,[b,c] Juliet A. Gerrard[d] and Craig A. Hutton[a,b,*]

[a]*School of Chemistry, University of Melbourne, Vic. 3010, Australia*
[b]*Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Vic. 3010, Australia*
[c]*Department of Biochemistry and Molecular Biology, University of Melbourne, Vic. 3010, Australia*
[d]*School of Biological Sciences, University of Canterbury, Christchurch, New Zealand*

Received 24 September 2007; revised 7 November 2007; accepted 8 November 2007
Available online 13 November 2007

Abstract—Dihydrodipicolinate synthase (DHDPS) is a key enzyme in lysine biosynthesis and an important antibiotic target. The specificity of a range of heterocyclic product analogues against DHDPS from three pathogenic species, *Bacillus anthracis*, *Mycobacterium tuberculosis* and methicillin-resistant *Staphylococcus aureus*, and the evolutionarily related *N*-acetylneuraminate lyase, has been determined. The results suggest that the development of species-specific inhibitors of DHDPS as potential antibacterials is achievable.

The development of narrow-spectrum antibacterials has several advantages, including limiting the development of drug-resistance[1] and minimising imbalances on important natural gut and intestinal flora.[2–4] While targeting genus-/species-specific proteins or membrane components has been one method used to achieve narrow-spectrum activity,[5] another is to target specific isoforms of essential bacterial enzymes.[6]

The bacterial diaminopimelate (DAP) pathway is responsible for the biosynthesis of the essential amino acid lysine and its immediate precursor *meso*-DAP, both of which are major constituents of the bacterial peptidoglycan cell wall.[7–9] Lysine is a constituent in Gram-positive bacteria (for example, the pathogenic bacterium *Staphylococcus aureus*) while the cell wall of Gram-negative bacteria, such as *Escherichia coli*, contains *meso*-DAP. Compounds that inhibit the DAP pathway may therefore represent a novel class of antibacterial agents. We have been engaged for some time in a study of the enzyme dihydrodipicolinate synthase (DHDPS, E.C. 4.2.1.52) which, as the first committed step of the biosynthetic pathway, is of particular interest as a drug target.[7,10–12] As part of this programme, we herein report on the potential for achieving species-specificity in the development of DHDPS inhibitors.

DHDPS is a homotetrameric enzyme that belongs to the *N*-acetylneuraminate lyase (NAL) sub-family of enzymes.[13] Each of the four monomeric units of DHDPS and NAL has a $(\beta/\alpha)_8$ barrel fold and these enzymes are believed to have evolved from a common ancestor.[14] The reactions catalyzed by DHDPS and NAL follow similar mechanisms, as shown in Figures 1 and 2.[13] Each enzyme has an active site lysine residue that condenses with pyruvate **1** to form a Schiff base. Tautomerization to the corresponding enamine is then followed by reaction with the relevant aldehyde-containing substrate (aspartate semi-aldehyde **2** or *N*-acetylmannosamine **6**) to generate a 4-hydroxy-2-iminoacid intermediate (**3** or **7**), which then proceeds to the heterocyclic product (HTPA **4** or sialic acid **8**). In the case of the DHDPS-catalyzed reaction, subsequent dehydration gives DHDP **5**.[15] Note that the NAL-catalyzed reaction is drawn in the reverse direction to the dominant physiological process, for comparison.[13]

Keywords: Dihydrodipicolinate synthase; DHDPS; Antibacterials; Species specificity; *N*-Acetylneuraminate lyase; NAL.
* Corresponding author. Tel.: +61 3 8344 6482; fax: +61 3 9347 5180; e-mail: chutton@unimelb.edu.au

0960-894X/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2007.11.026

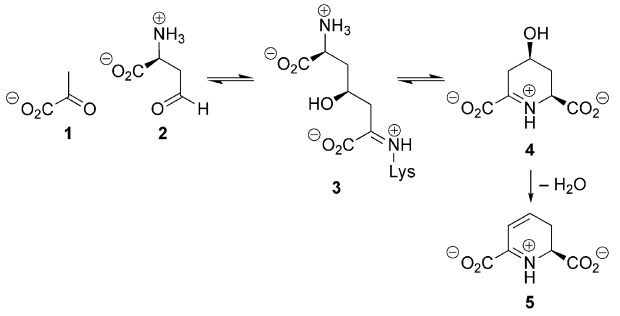


Figure 1. DHDPS-catalyzed reaction.

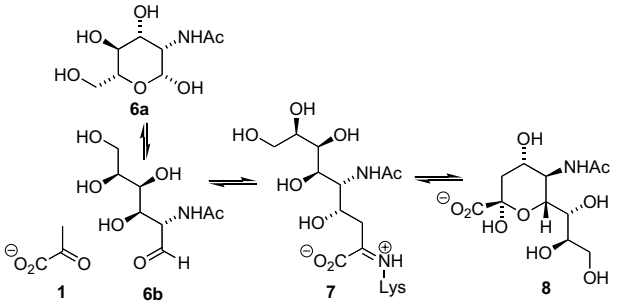


Figure 2. NAL-catalyzed reaction (reverse of physiological direction).

In earlier work, analogues of the DHDPS product 4-hydroxytetrahydrodipicolinate (HTPA **4**) were screened against *E. coli* DHDPS and shown to be moderate, non-competitive inhibitors.[10] Accordingly, we have investigated the efficacy of these compounds against DHDPS from a range of pathogenic bacterial species, in order to establish whether species selectivity of such compounds is feasible. The compounds were also screened against *E. coli* NAL, in order to establish whether the binding site for these inhibitors has been conserved across related enzyme families.

DHDPS from three pathogenic bacterial species, *Bacillus anthracis*,[16] *Mycobacterium tuberculosis*[17] and methicillin-resistant *S. aureus* (MRSA),[18] and *E. coli* NAL[13,18] were expressed in and purified from *E. coli*.

Compounds **9**–**18** (Fig. 3) were synthesised according to the reported methods[10] and were screened against all enzymes (Table 1). In the DHDPS inhibitor screen, our standard coupled assay was employed, in which the coupling enzyme was DHDPR.[12,19,20] For screening NAL, an analogous assay was used employing lactate dehydrogenase as a coupling enzyme[21,22].

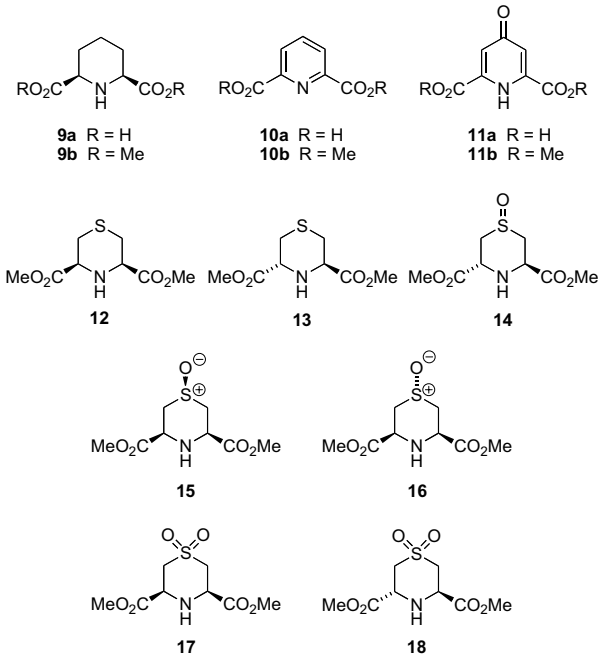


Figure 3. Heterocyclic compounds **9**–**18** tested for inhibition of DHDPS and NAL.

Table 1. Inhibition assays against bacterial DHDPS and NAL[a]

Compound	Inhibition of enzyme activity[b]				
	E. coli DHDPS	*B. anthracis* DHDPS	*M. tuberculosis* DHDPS	MRSA DHDPS	*E. coli* NAL
9a	49	73	43	51	21
9b	92	99	24	58	18
10a	76[c]	80	75	67	1
10b	5	13	0	0	0
11a	74[d]	0	73	83	27
11b	85[e]	0	84	88	64
12	35	23	0	0	1
13	12	19	0	0	0
14	0	7	0	7	7
15	20	22	14	12	38
16	8	8	1	10	0
17	0	19	6	0	3
18	14	6	7	0	0

[a] Assays were performed in duplicate and were typically within ±3%.
[b] % Inhibition in the presence of 20 mM **9**–**18**.
[c] K_i = 11 mM versus **1**, 18 mM versus **2** (Ref. 20).
[d] K_i = 22 mM versus **1**, 25 mM versus **2** (Ref. 10).
[e] K_i = 7 mM versus **1**, 14 mM versus **2** (Ref. 10).

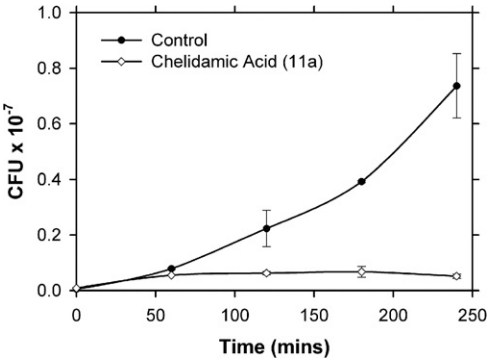


Figure 4. Inhibition of *E. coli* growth by chelidamic acid **11a**.

The DHDPS assays show that piperidine-2,6-dicarboxylate **9a** and dipicolinic acid **10a** are relatively potent inhibitors across all species with little selectivity.

In a dramatic example of species selectivity, chelidamic acid **11a**, and its diester **11b**, are potent inhibitors of DHDPS from *E. coli*, *M. tuberculosis* and MRSA, but in contrast display no inhibition at all of *B. anthracis* DHDPS. Similarly, several compounds, including piperidine diester **9b** and the thiazanes **12** and **13**, exhibit significantly greater inhibitory activity against *E. coli* and *B. anthracis* DHDPS than DHDPS from *M. tuberculosis* and MRSA.

Few of the compounds studied exhibited significant inhibition of *E. coli* NAL, with dimethyl chelidamate **11b** being the only one to show >50% inhibition at the concentration tested, and thiazane-*S*-oxide **15**, piperidine-diacid **9a** and chelidamic acid **11a** exhibiting lower activity. That *E. coli* NAL was inhibited to some degree by several of the compounds suggests that the inhibition site has been conserved across families and may therefore have some functional significance.

In order to validate the inhibitors as potential leads for the development of antibacterial agents, the antibacterial activity of selected compounds against *E. coli* was determined. Compounds **9b**, **11a** and **17** were chosen for analysis: piperidine diester **9b** as it exhibited the most potent inhibition of *E. coli* DHDPS, chelidamic acid **11a** as it shows significant species selectivity and sulfone **17** as a control as it displayed very low levels of DHDPS inhibition across all species tested. Chelidamic acid **11a** at a concentration of 20 mM displayed strong inhibition of bacterial growth (Fig. 4) relative to the control culture in the absence of **11a**. Piperidine ester **9b** also at a concentration of 20 mM displayed moderate inhibition of growth, whereas sulfone **17** displayed virtually no inhibition. These results suggest that DHDPS inhibitors are able to confer significant antimicrobial activity.

In conclusion, several compounds displayed clear differentiation in inhibition of DHDPS enzymes from different bacterial species, which suggests that latter generation compounds could be targeted to specific pathogens. Validation of the potential of these compounds as leads for the development of antibacterials was demonstrated; for example, chelidamic acid **11a** displayed high levels of inhibition of both DHDPS activity and bacterial growth, in contrast to sulfone **17**, which displayed no enzyme inhibition and no antibacterial activity.

Acknowledgments

This work was funded in part by the Australian Research Council (LX0776388 and DP0770888) and in part by the Royal Society of New Zealand Marsden Fund (contract UOC303).

References and notes

1. Lee, V. J.; Miller, G. H.; Yagisawa, M. *Curr. Opin. Microbiol.* **1999**, *2*, 475.
2. Myllyluoma, E.; Ahlroos, T.; Veijola, L.; Rautelin, H.; Tynkkynen, S.; Korpela, R. *Int. J. Antimicrob. Agents* **2007**, *29*, 66.
3. Nord, C. E.; Kager, L.; Heimdahl, A. *Am. J. Med.* **1984**, *76*, 99.
4. Norrby, S. R. *Scand. J. Infect. Dis. Suppl.* **1986**, *49*, 189.
5. Fernandes, P. *Nat. Biotechnol.* **2006**, *24*, 1497.
6. Kotaka, M.; Dhaliwal, B.; Ren, J.; Nichols, C. E.; Angell, R.; Lockyer, M.; Hawkins, A. R.; Stammers, D. K. *Protein Sci.* **2006**, *15*, 774.
7. Hutton, C. A.; Perugini, M. A.; Gerrard, J. A. *Mol. BioSyst.* **2007**, *3*, 458.
8. Hutton, C. A.; Southwood, T. J.; Turner, J. J. *Mini-Rev. Med. Chem.* **2003**, *3*, 115.
9. Cox, R. J.; Sutherland, A.; Vederas, J. C. *Bioorg. Med. Chem.* **2000**, *8*, 843.
10. Turner, J. J.; Gerrard, J. A.; Hutton, C. A. *Bioorg. Med. Chem.* **2005**, *13*, 2133.
11. Turner, J. J.; Healy, J. P.; Dobson, R. C. J.; Gerrard, J. A.; Hutton, C. A. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 995.
12. Coulter, C. V.; Gerrard, J. A.; Kraunsoe, J. A. E.; Pratt, A. J. *Pesticide Sci.* **1999**, *55*, 887.
13. Lawrence, M. C.; Barbosa, J. A.; Smith, B. J.; Hall, N. E.; Pilling, P. A.; Ooi, H. C.; Marcuccio, S. M. *J. Mol. Biol.* **1997**, *266*, 381.
14. Joerger, A. C.; Mayer, S.; Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 5694.
15. Blickling, S.; Renner, C.; Laber, B.; Pohlenz, H. D.; Holak, T. A.; Huber, R. *Biochemistry* **1997**, *36*, 24.
16. Blagova, E.; Levdikov, V.; Milioti, N.; Fogg, M. J.; Kallioma, A. K.; Brannigan, J. A.; Wilson, K. S.; Wilkinson, A. J. *Proteins* **2006**, *62*, 297.
17. Kefala, G.; Weiss, M. S. *Acta Crystallogr., Sect. F* **2006**, *62*, 1116.
18. Details to be published elsewhere.
19. Yugari, Y.; Gilvarg, C. *J. Biol. Chem.* **1965**, *240*, 4710.
20. Karsten, W. E. *Biochemistry* **1997**, *36*, 1730.
21. Comb, D. G.; Roseman, S. *Methods Enzymol.* **1962**, *5*, 391.
22. Aisaka, K.; Uwajima, T. *Appl. Environ. Microbiol.* **1986**, *51*, 562.